计算机视觉
——基于图像的3D重构

于永彦◎著



说　明

本书版权属于北京大学出版社有限公司。版权所有，侵权必究。

本书电子版仅提供给高校任课教师使用，如有任课教师需要本书课件或其他相关教学资料，请联系北京大学出版社客服，微信手机同号：15600139606，扫下面二维码可直接联系。

由于教材版权所限，仅限任课教师索取，谢谢！



计算机视觉
——基于图像的 3D 重构

于永彦　著

北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

本书在介绍计算机视觉基本知识的基础上，重点探讨 3D 重构的各种理论及其内在联系，以及经典的应用案例，并对其中的相机标定、多模型估计等关键技术提出独到的见解，规划一整套算法，构建完善的学科体系。

图书在版编目(CIP)数据

计算机视觉：基于图像的 3D 重构 / 于永彦著. —北京：北京大学出版社，2020.12

ISBN 978-7-301-31859-1

Ⅰ. ①计… Ⅱ. ①于… Ⅲ. ①计算机视觉 Ⅳ. ①TP302.7

中国版本图书馆 CIP 数据核字（2020）第 230053 号

书　　名	计算机视觉——基于图像的 3D 重构 JISUANJI SHIJUE——JIYU TUXIANG DE 3D CHONGGOU
著作责任者	于永彦　著
策划编辑	郑　双
责任编辑	郑　双
标准书号	ISBN 978-7-301-31859-1
出版发行	北京大学出版社
地　　址	北京市海淀区成府路 205 号　100871
网　　址	http://www.pup.cn　新浪微博：@北京大学出版社
电子信箱	pup_6@163.com
电　　话	邮购部 010-62752015　发行部 010-62750672 编辑部 010-62750667
印 刷 者	
经 销 者	新华书店 650 毫米×980 毫米　16 开本　12.75 印张　212 千字 2020 年 12 月第 1 版　2020 年 12 月第 1 次印刷
定　　价	68.00 元

前　言

基于图像的3D重构是计算机视觉领域的研究热点，被广泛应用于航空航天、机械制造、医疗、考古、地质、犯罪现场复原、建筑设计、城市规划等领域。相对于较成熟的多视图或多幅图像的3D重构技术，基于单幅图像的3D表面重构由于其先天性约束不足和潜在的巨大价值受到各国研究人员前所未有的关注。本书针对现有单图像3D重构技术中的SFS、SFT等两种典型方案的局限性，分别提出有针对性的改进措施，并提出一种新的基于图像中几何元素共平面性约束的3D表面重构方案。

第一，针对SFS类方法大多严重依赖边界条件和初始假定解而可能导致多重解、无解或局部解的歧义性问题，提出基于表面可积性条件将SFS求解目标规划为一个仅含二次项约束的多项式系统，使用同伦分析方法提高解的存在概率和解的精确度；基于精确线搜索技术，通过设计新的平滑项优化迭代过程，加快收敛速度；针对多项式系统可能存在非凸项问题，构造对应的SDP凸松弛，使SFS摆脱了对初始假定解的严重依赖性，确保SFS迭代收敛于一个全局极小值，从而获得闭合解。实验结果表明，这种方案大大提高了SFS求解过程的高效性和结果的普适性，而且对弹性表面和非弹性表面都有较好的重构效果。

第二，将纹理表面中普遍存在的局部面片法向倒置等二进制歧义性归结为一种离散判定问题，首先将表面表示为一个张量积基函数的线性组合，用离散变量描述该歧义性，进而定义一个二次目标函数来度量表面平滑度和约束满足度，通过消除连续变量产生一个仅含离散变量的二次目标函数，再使用SDP凸松弛法，将离散变量嵌入一个连续高维空间，通过基于分裂惯性和角扫描的圆整操作将离散变量归约到一个低维空间，最后通过最大割算法获得较理想的解。上述方案比传统的椭球法高效且成本低，还可进一步使用K-L算法改善局部搜索过程，消除错误割点，有效提高解的精确度。

第三，针对图像中几何元素的共平面性可提供景深信息的特点，提出一种基于交叉平坦曲线共平面约束的新的3D表面重构方案，即

对于不平行于投影方向的某一个平面，根据其所含曲线与另一平面中某曲线的交叉构型构造一个线性系统，当一组这样的交叉曲线位于拟求解表面时，可获得该系统的精确解空间。对于含噪系统，增加新的约束条件，使用 SVD 法可得到逼近解。再利用正投影和透视投影的等价性，可将透视投影转化为正投影，从而将这两种投影下的 3D 表面重构归入到一个框架中。实验表明，这种方法大大提高了单图像 3D 表面重构的健壮性，可适用于完全未标定结构光等真实场景。

本书在内容安排上按照由浅入深、由简到繁、由基础点到专业体系的基本思路来组织素材，力求在理论讲解上既不空洞又有趣味，在技术实现上要求可操作型强，让读者一看就懂、一试就成功，目的是激发读者的尝试欲望，培养、提高读者的实践能力。

本书的组织结构如图 0-1 所示(其中数字代表章节序号)。具体内容介绍如下。

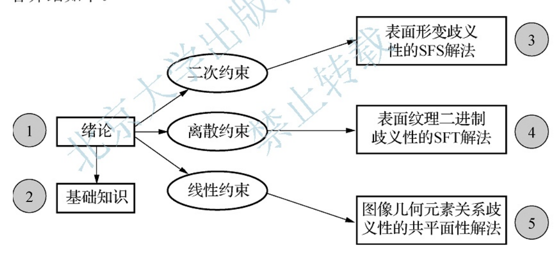


图 0-1　本书的组织结构图

第 1 章，首先介绍计算机视觉的基本概念，接着讨论基于图像的 3D 表面重构的发展概况及数学原理，重点分析基于单幅图像的 3D 表面重构技术的非适定现象，最后简要总结了各种歧义性问题，并由此提出 3 种基本的约束模式。

第 2 章，扼要介绍 3D 表面的形式定义及成像模型，着重分析平滑表面成为局部平面的必要条件，并简单讨论 Lambertian 表面的假设条件，以及单光源假设的合理性和必要性。最后系统介绍同伦分析法

(Homotopy Analysis Methed，HAM)和半定规划(Semi-Definite Programming，SDP)两种较流行的优化方法。

第 3 章，介绍 SFS 方法的理论基础及存在的问题，提出将 SFS 问题规划为多项式结构，再讨论如何将成熟的同伦分析法、优化线搜索法和半定规划法等用于求解多项式。

第 4 章，介绍表面纹理的二进制歧义性，提出将含二进制歧义性的 SFT 归结为一个离散判定问题，再通过半定规划凸松弛、圆整操作和最大割算法等步骤获得较理想的逼近解。

第 5 章，介绍图像中几何元素的共平面性特征，并提出一种基于交叉平坦曲线的共平面约束的 3D 表面重构方法。

第 6 章，总结本书所做的工作，对主要创新点分别做了小结，并对未来研究方向提出了展望。

本书在写作过程中得到了王志坚教授、娄渊胜教授、吕鑫博士以及其他很多热心人士的帮助与支持，感谢内江师范学院的领导及各位同人，感谢他们为全书的整体性构思提供了许多建设性的建议，感谢他们为本书提供了实验平台，还有其他为本书进行过文字校对、排版的老师及同学，此处不再一一列出，在此一并表示最诚挚的谢意！

虽然作者主观上做了最大的努力，但由于本身的水平有限，加上时间仓促，难免存在一些不足。“他山之石，可以攻玉”，真诚地希望使用或阅读本书的读者给予批评指正，不吝赐教。

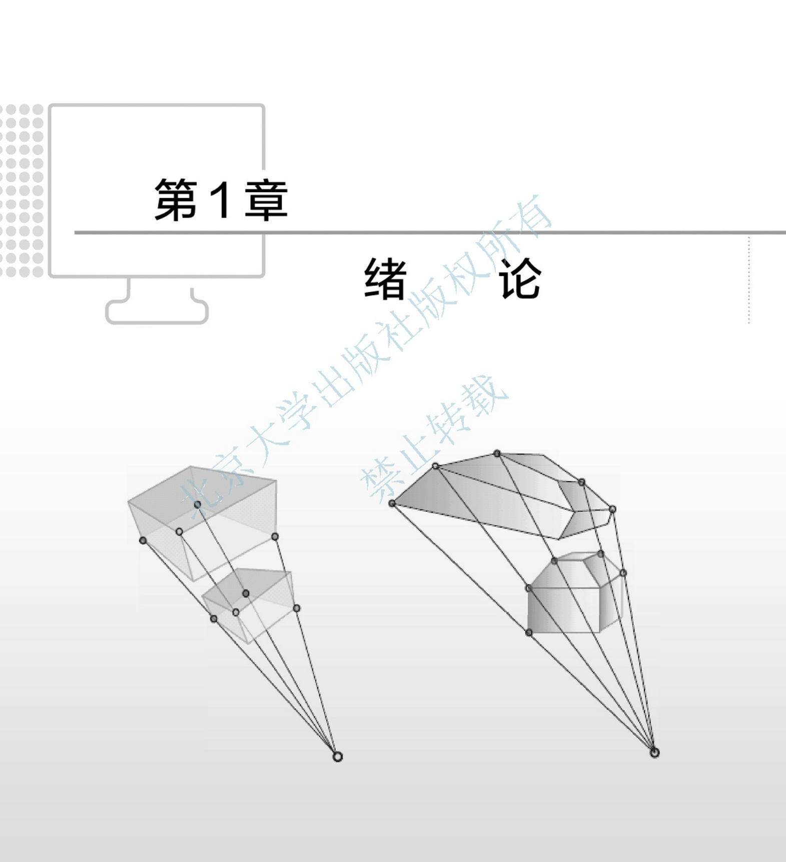
于永彦

2020 年 7 月 3 日　于 四川内江

目　　录

第1章

绪　论



伴随着 20 世纪 40 年代第一台电子计算机的横空出世，人们就一直在探讨一个经典的哲学话题，即人类“能否制造出一台像人一样聪明的智能机器”，这是近一个世纪以来激励计算机科学蓬勃发展的核心动力之一，直接引领了人工智能学科的诞生。人工智能的一个重要分支称为计算机视觉，主要研究如何使计算机具备“看”的能力，在生产生活的各个领域得到了广泛应用，如无人驾驶的导航系统、工业上的非接触测量、医疗领域的图像辅助诊疗、智能交通、安全监控系统及家庭娱乐等，并由此极大地促进了相关学科的协同发展，如视觉生理学、心理学、几何学、统计学、物理学、信息论、拓扑学、视觉计算理论、优化理论等。

1.1 引　言

视觉系统是人类生理系统中构造最复杂、功能最强大的感官系统。使用计算机来模仿人类视觉，就称为计算机视觉。计算机视觉技术专门研究如何协助机器来模拟人类的视觉感知能力，可以像人类那样准确观察、分析、理解和再现真实世界的各类场景结构。图 1.1 描述了计算机视觉的基本原理与一般过程。

计算机视觉的数学基础是投影几何，本书主要涉及正投影和透视投影。正投影也称平行投影，所形成的图像较为直观，数学模型简单，易于理解。而透视投影可看成正投影的一般形式，或者说正投影是透视投影的特例。透视投影是指投影中心位于理想的无穷远点处。

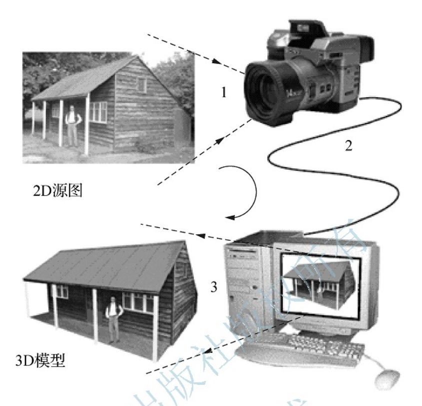


图 1.1　计算机视觉的基本原理与一般过程

图 1.2 描述了最简单的针孔相机的投影成像模型。三维空间中某点 M 在二维平面上的投影点 m 可表示为式(1.1)。

$$m \propto \boldsymbol{F}\boldsymbol{M}_c \tag{1.1}$$

式(1.1)中的 $\boldsymbol{F}$ 表示针孔成像的投影矩阵。

计算机视觉的发展大致经历了两个阶段。第一个阶段为计算视觉理论。英国心理学家 Marr[1]在对人类视觉系统全面研究的基础上，将心理学、人工智能和神经生理学的已有成果结合起来，于 1982 年首次提出较为完善的计算机视觉理论的基本框架[2](见图 1.3)，促进了计算机视觉学科的形成，为计算机视觉的蓬勃发展与工程化应用奠定了重要的理论基础。Marr 从信息处理系统的角度出发，将视觉系统的研究分为三个层次，即计算理论层次、表达与算法层次、硬件实现层次。

计算理论层次主要回答系统各部分的计算目的与计算策略；表达与算法层次给出各部分的输入输出和内部的信息表达，以及实现计算理论所规定的目标的算法；而硬件实现层次则回答“如何用硬件实现以上算法”。

图 1.2　针孔相机成像模型

低层阶段：在图像坐标系中，对原始2D图像进行处理，抽取图像中的基本图元，如拐点、边缘、纹理、线条、边界等

中层阶段：在相机坐标系中，由原始图像和基本图元恢复场景中可见部分的深度、法向、轮廓等，构成2.5D图

高层阶段：在世界坐标系中，由输入图像、基本图元、2.5D图来恢复、表示、识别3D物体

图 1.3　计算视觉理论的基本框架

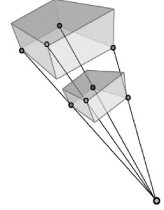
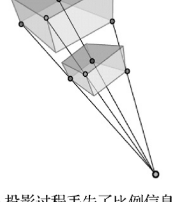
Marr 计算视觉理论无疑是计算机视觉研究领域的划时代成就，指出了研究计算机视觉许多珍贵的哲学思想和研究方法，同时也给计算机视觉研究领域开创了许多研究起点。但由于时代的局限性，该理论不可避免地存在一些有争议的方面。例如，其提出的计算框架基本上是单向的，没有反馈，以及对先验知识的应用没有引起足够重视等。

第二个阶段是在 20 世纪末发展起来的“几何计算”视觉理论。它把复杂的数学，尤其是射影几何引入到计算机视觉的研究中，形成“多视图几何”，为理解和形式化多视图成像几何奠定了坚实的理论基础，使得以前被认为无法求解或难以求解的问题得以顺利解决。

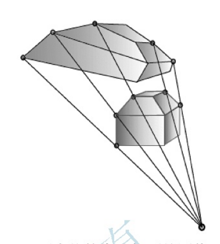
需要指出的是，无论在哪个阶段，基于图像的 3D 重构都是计算机视觉的一个至关重要的核心领域和热点问题，也是用计算机模拟人眼视觉功能所需要完成的关键步骤。

1.2　基于图像的 3D 重构

计算机视觉的终极目标是从所“看”到的场景中识别出感兴趣的对象，即从 2D 图像中重构出 3D 对象模型，主要包含图像处理、模式识别、景物分析、图像理解等基本处理过程。但是，由于投影过程中丢失了景深信息或因遮挡而遗漏了部分有用信息，将可能导致不同形状的 3D 对象投影到图像平面上形成了相同或类似的图像，这就是投影歧义性。图 1.4 展示了常见的投影歧义性。

(a) 投影过程丢失了比例信息



(b) 形变物体产生相同的图像

图 1.4　常见的投影歧义性

另外，其他因素的不良影响，诸如光线差异、观察角度偏差、研究人员的认知水平与经验存在较大差距等，都可能给图像重构造成极大的困难。所有这些都可能导致基于图像的 3D 重构陷入一种病态。

在计算机视觉领域，基于图像的 3D 重构一直是被普遍关注的研究热点，也是大多数基于计算机的事件模拟和可视化应用的普遍需求，包括地学分析、机器人导航、犯罪场景复原、飞行模拟、游戏娱乐、特技表演、建筑设计、旅游指南及虚拟博物馆等。另外，3D 重构技术在考古及古迹保护方面也有特殊的应用。由于不可预测的火灾、地震、洪水、自然腐蚀、战争等天灾人祸，大量的历史文物面临抢救性保护，需要为子孙后代或历史研究保留一些珍贵遗迹，或者复原考古发现的珍贵文物，这些都需要借助 3D 重构技术来修复、重建。另外，3D 电视的诞生和普及，也需要通过 3D 重构技术将现有的大量 2D 资源转换为 3D 视频。所有这些，对于人类视觉系统而言也许是轻而易举的事，但对于机器视觉而言，则面临巨大的挑战，既需要突破现有的理论瓶颈，又需要寻找合理可行的实践方法。

1.2.1　3D 重构的数学原理

人眼观察到的都是 3D 现实世界。使用数码相机等普通设备对周围场景拍照或摄像，得到的一般是 2D 真实图像；或者使用 3D 扫描设备采集 3D 场景数据，再通过图像重建得到合成图像。这个过程统称为成像。图 1.5 演示了图像合成的数学原理。

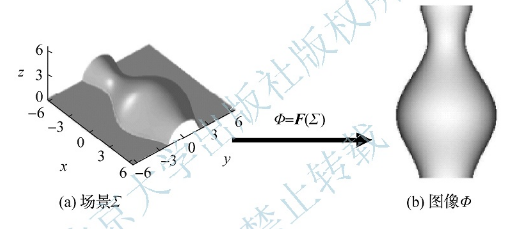


图 1.5　图像合成的数学原理

基于图像的 3D 重构可看成上述成像过程的逆过程，是指根据 2D 图像再现 3D 物体形状。这两个过程中的数学原理见表 1-1 和表 1-2。

表 1-1　成像过程及其逆过程的数学原理

问题	已知量	未知量	技术
正向过程	Σ (真实场景或 3D 数据)， $\boldsymbol{F}$ (投影矩阵)	Φ	图像合成、摄像等
逆向过程	Φ	Σ，$\boldsymbol{F}$	3D 重构

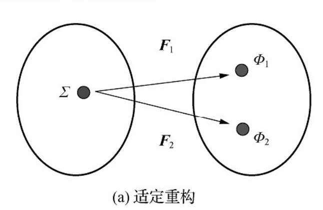
表 1-2　适定重构与非适定重构的数学原理

问题	已知量	未知量	技术
适定重构	$(\varPhi_1,\boldsymbol{F}_1)$，$(\varPhi_2,\boldsymbol{F}_2)$	$\varSigma$	立体视觉
非适定重构	$(\varPhi,\boldsymbol{F})$	$\varSigma_1$，$\varSigma_2$	单目视觉

根据已知量的组成结构，基于图像的 3D 重构技术大致可分为适定和非适定两种类型。

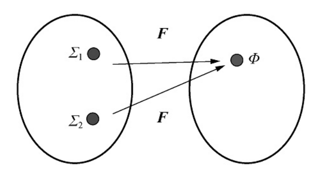
(1) 适定重构。即满足解是存在的、解是唯一的、解连续依赖于定解条件三个基本条件。图 1.6(a)表示了同一场景可以对应两幅以上不同的图像。这种情况下的 3D 重构理论和技术目前较为成熟，较典型的是立体视觉重构，得到了较广泛的应用。

(2) 非适定重构。即上述三个条件至少有一条不满足。图 1.6(b)表示可能存在两种不同的场景对应同一幅图像。此时由于投影过程的信息丢失，致使重构所需信息严重缺乏，重构过程存在极大的偶然性，重构结果不唯一，需要在众多有歧义的解簇中确定唯一解。非适定重构目前是研究的热点，也是难点。



(a) 适定重构

图 1.6　适定与非适定 3D 重构



(b) 非适定重构

图 1.6 适定与非适定 3D 重构(续)

适定重构与非适定重构的数学原理可用表 1-2 描述。

由上述可知，基于适定重构的立体视觉可以很好地避免 3D 重构的病态现象。例如，对于同一个场景，采用两个相机，从不同角度得到多幅不同视差的图像，从而可得到场景的深度信息。但是，人类视觉却可从单幅图像中轻松获得深度信息，因为人们可以利用日常积累的经验知识。因此，基于单幅图像的 3D 重构一直是计算机视觉领域极具挑战性的研究热点之一，在逆向工程、医学图像辅助治疗、建筑学、考古及古迹保护、地质考察、遗迹复原等领域具有巨大的应用价值。

1.2.2 发展概况

由式(1.1)可知，基于图像的 3D 重构的根本任务是，根据 2D 图像点 m 求出对应的空间点 M，进而估算 3D 场景及所包含对象的深度信息，最终解决“图像看起来像什么”的问题。其本质是从 2D 图像观察 3D 几何信息，其关键步骤是精确标定投影矩阵 $\boldsymbol{F}$。

1. 基本理论

目前较常用的 3D 重构理论是分层重构理论，即分三个层次分别

实现对象重构。第一个层次是射影重构，即从图像获得 3D 物体的射影结构。此阶段不需要附加任何先验知识，仅从图像本身即可获得所需信息；第二个层次是仿射重构，即利用先验知识获得物体的仿射结构，需要知道无穷远平面或无穷远单应性矩阵；第三个层次是度量重构，即利用先验知识获得物体的度量结构，一般需要知道相机的内外参数，或绝对二次曲线、绝对二次曲面等可用于标定相机参数的图像测量信息。

经过数十年的发展，目前已涌现出许多优秀的 3D 重构理论、技术和算法，有针对图形的，如机械制造中基于三视图的重构，主要目标是保证重构结果的真实可靠和理解上的无二义性；有针对图像的，如遥感图像、CT 图像等；有基于多视图的[3~6]，也有基于单视图的[7~11]。

2. 基本方法

针对基于图像 3D 重构过程的主体对象，现有重构方法大致可分为两大类。一类是主动式，即利用场景或对象自身的一些特征，如几何特征和纹理特征等来提取物体表面的 3D 信息并进行建模。主动式重构对工作环境、摄像设备没有特殊要求，容易实现，在军事、遥感测量、机器人导航等方面有着广泛的应用基础，受到各国研究人员的高度重视，取得了丰硕的成果，涌现出一大批有效的算法。其中最著名的是 SFM(Shape From Motion，基于运动的形状表面重构)[12]，并由此衍生出一大类算法，如 SFPS(Shape From Photometric Stereo，基于光度立体视觉的形状表面重构)、SFS(Shape From Shading，基于明暗度的形状表面重构)、SFT(Shape From Texture，基于纹理的形状表面重构)、SFC(Shape From Contour，基于轮廓的形状表面重构)等。这些

方法的一个共同缺陷是，当用于单幅图像的 3D 重构时表现为病态解，因而需要对场景做一些额外假定，并且需要具备重构对象的某些先验知识，当理想条件与实际情况有较大出入时，重构结果可能变得面目全非，而且难以获得高精度的 3D 信息。另一类是被动式[13~16]，即需要手工在物体表面做某种标记，如条纹或阴影，并且根据这些信息建模或由传感器获得三维信息。因此，这种方式需要使用专门的扫描光源等比较复杂的设备，故获取的 3D 信息精度较高。

目前主要是利用激光扫描的各种结构光方法，但其设备比较复杂，体积、重量比较大，造价也十分昂贵，其作用距离也比较短。目前的大多数技术都需要分析提取真实世界的 2D 几何特性，如垂直面[17]、水平面[18]、矩形结构[19]、对称结构[20]等。对于单幅图像通常需要交互处理，即由用户提供场景的结构信息和约束条件，如消隐点、消隐线、共面等。交互 3D 重构[13,14,21,22,23,24,25,26]不需要相机的任何先验知识，可在重构过程中根据获得的信息估计相机的各类参数。

若根据 3D 重构对象的不同，也可大致分为两类：一是照片几何方法，即基于照片中几何对象的模型化；二是光度测量法，主要基于光度测量模型化。不同的具体应用背景，对所需输入信息的需求也不一样。有的重构任务仅需要 3D 空间点的稀疏集即可，即只需要输入图像的少部分像素信息。而有的重构任务则需要稠密重构，需要考察输入图像的每一个像素的尽可能多的信息。另外还可能需要其他一些重要的信息，诸如物体对象的尺寸、形状及材质，输入图像的数量，光源或相机的相对位置等。

3. 基本问题

关于 3D 重构，主要涉及三个问题。

(1) 输入数据的数量或类型。这依赖于输入视图的多少。多视图重构[3,4,5,6]使用场景或对象的两个以上的视图，而单视图[7]仅使用单个视图。立体视图[27]是多视图的特例，使用两个视图。

(2) 输出的一般化描述。主要有容积表示[11]和表面表示[28]。容积表示将包围物体的体积离散化为体元，从而估计可以按某种度量规则拟合输入图像的体元的占有率。表面表示法为物体构造虚拟的网状覆盖面[29]，即可观察的外部轮廓。一般来说，多视图重构采用一般化表示法，较少使用先验知识。

(3) 先验知识的复杂性。为了保证更多的对象能被精确估计，通常定义概率密度(即能量函数)作为一种先验知识，可用于定义两个重构结果之间的微小差异。例如，容积表示中的平滑性[30]、削面的平坦性[31]、角度的相似性[32]、基于表面重构的长度[10]、能量函数的平滑性[8]、连接角的生存发育性[9]等。相比于强制约束，定义先验知识的好处是，可以针对特定数据自动训练出先验知识，而不必重新设计新的表示方法。

正如前面所述，基于图像的 3D 重构是一个病态逆向问题，如果没有额外的假定条件，基于单幅图像的重构可能存在无数解，这是因为存在歧义性或存在重叠图案。因此，需要附加某些假定条件以消除或缓解歧义性。但是，这些假定通常并非总是有效，尤其在实际世界中，如树、丛林等。另外，即使消除了歧义性，获得精确解仍然是个难题。因为在成像过程中不可避免地存在各种噪声和畸变失真，这些现象都难以被模型化或有效补偿。因此，根据具体的场景构型寻找一种有效、健壮的数学模型是 3D 重构的核心问题。例如，若已知研究对象是一辆汽车，则可使用特殊模型(对称性、轮胎椭圆性等)来约束上述重构模型。其他的，诸如城市街道、建筑物、天体图像等都可按这种思路进行。

1.3 基于单幅图像的 3D 重构

基于单幅图像的 3D 重构是计算机视觉领域最具挑战性的研究热点，由于其所能提供的信息非常有限，因此是一个经典的非适定病态难题。但是，对于人类视觉系统来说，可以借助一些关于所研究对象的先验知识或日常经验常识来从单幅 2D 图像恢复 3D 模型。因此，很多研究人员根据这个原理，提出了很多假设条件，如几何形状的平坦性[13]、平行性及垂直性[33]、表面纹理的畸变性[23]、局部表面平滑性[34]、明暗渐变性[14]、散焦或聚焦性[25]、人机交互方法[35]等，试图根据单幅图像再现空间场景。这些方法统称为有约束重构问题。它们的缺点是，这些先验知识一般是试探性的，仅适合于特殊场合或特定领域问题，如人脸[36]、建筑物[28]、医学图像[37]等，难以推广到一般模型。

从数学原理角度看，基于单幅图像的 3D 重构主要依据图像中包含的深度信息 $z(x,y)$、表面法向 $\boldsymbol{N}(x,y,z)$、表面梯度 (p,q)，以及表面的倾角和斜角等信息恢复表面形状。图 1.7 是一个典型的例子。

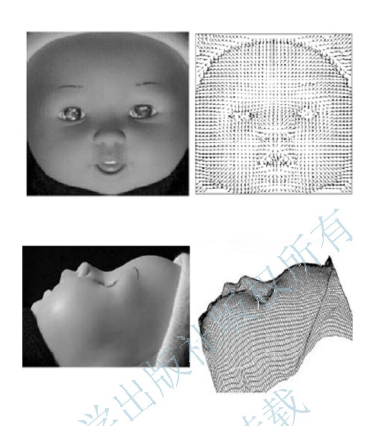


图 1.7　根据表面法向重构表面形状

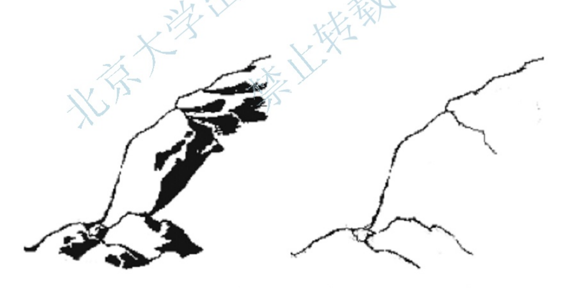
上面所述的深度信息可以是相机和对象表面点之间的距离，也可以是表面与参考平面 $X-Y$ 之间的相对距离。表面法向是指垂直于对象表面某点的切平面的法向。表面梯度 $(p,q)=(\partial z/\partial x, \partial z/\partial y)$ 是指深度信息在 x, y 方向上的变化率。倾角 ϕ 和斜角 θ 是指表面点相对于法向的角度，可表示为 $(x,y,z)=(l\sin\phi\cos\theta, l\sin\phi\sin\theta, l\cos\phi)$，其中 l 是表面法向的一个度量。在计算机视觉中，从图像恢复形状的技术统称为 SFX(Shape From X)，其中的 X 可以是明暗度(shading)、纹理(texture)、立体视觉(stereo)、运动(motion)等。目前较有影响的是 SFS 和 SFT，下面分别做一些探讨。

1.3.1　SFS

SFS 是计算机视觉中的经典研究热点问题，主要根据图像中明暗(阴影)渐变性来恢复表面法向和表面形状(深度)。其基本步骤是：首先将图像的像素光亮度(有时称为光强)模型转化为表面几何的函数，然后恢复一个表面，再生成一个图像，以逼近给定的原图像。

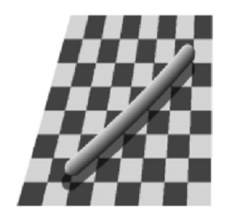
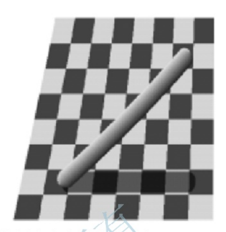
1. 阴影几何基础

艺术家们很早就注意到阴影在 3D 绘图中的重要性，可用来表示场景的表面形状或空间位移。图 1.8 展示的虽然是 2D 图像，但仍然表现出强烈的 3D 结构。



(a) 内阴影：提供形状感知的凸角信息和场景照明方向信息

图 1.8　阴影在 3D 表面感知中的重要作用

(b) 外阴影：提供对象的相对位置和方向的线索信息

图 1.8　阴影在 3D 表面感知中的重要作用(续)

(1) 内阴影由物体对象在自身形成。沿着内阴影边界的点提供了表面形状约束，可作为表面感知的边界约束。内阴影包括两种类型：attached 阴影和 cast 阴影。attached 阴影区域包含了法向方向背离光源方向的表面上的点，而 cast 阴影区域包含的点面向光源，但被表面的一个末梢部分遮挡，也就是说，cast 阴影区域包含了表面上被遮挡的点，如图 1.9 所示。

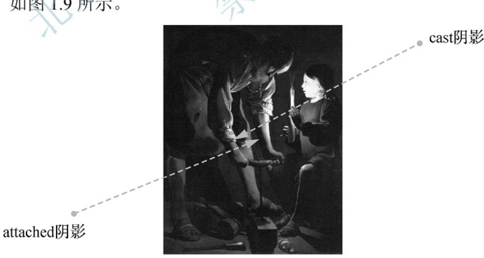


图 1.9　attached 阴影和 cast 阴影

(2) 外阴影由一个发光的物体投射光线到另一个物体上而形成。外阴影轮廓提供了表面的几何信息，而内阴影提供了表面上明暗图案的一个粗糙的感知结果。

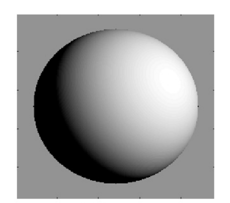
不过，由阴影提供的信息性质上不同于由明暗提供的信息，前者本质上包含在由阴影边界、表面、光照几何等构成的曲线之间的几何关系中。而且，自然观察条件下的阴影并非仅是等照度线的一种特例。这里所说的等照度线是指图像中亮度恒等的曲线集合，由多个因素决定，包括表面的内部反射及表面间的交互反射等。

阴影包含了图像分割的多种线索，至少在理论上可以被准确度量。人们研究阴影，目的是构造 SFS 方法，可用于解决计算机视觉领域的多个问题。例如，沿着内阴影边界的点提供了表面形状的约束，可用于表面插值的边界条件。在阴影内部区域，如被遮挡的轮廓，可用于约束 SFS。

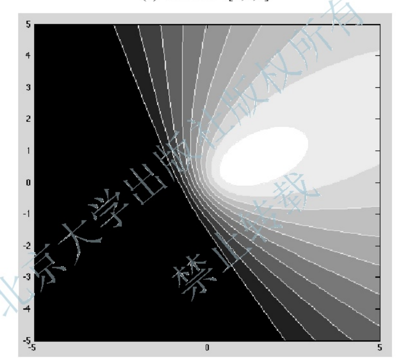
基于阴影边界的主要约束是：①沿着一个 attached 阴影轮廓的点处的表面法向垂直于光源方向；②图像中位于同一光束的 attached 和 cast 阴影中的点，位于 3D 中同一光束上。虽然 SFS 问题是显著约束不足的，但人类却可以感知出唯一的、稳定的形状。

感知 SFS 的理论线索主要有以下几个方面。

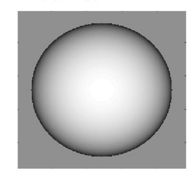
(1) 反射图。将场景图像的光亮度表示为 3D 表面方向的一个函数，忽略阴影、内部交互反射、渐晕、半透明等现象，如图 1.10 所示。



(a) 光照方向[0,0,1]

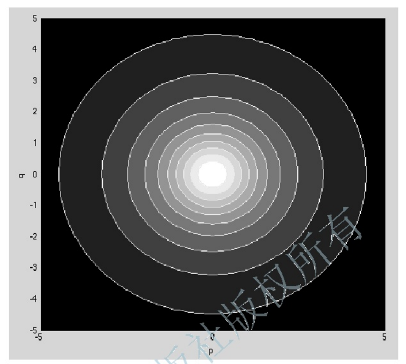


(b) 图(a)的反射图



(c) 光照方向[1,0.5,1]

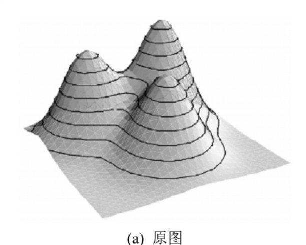
图 1.10　反射图与光照方向的关系



(d) 图(c)的反射图

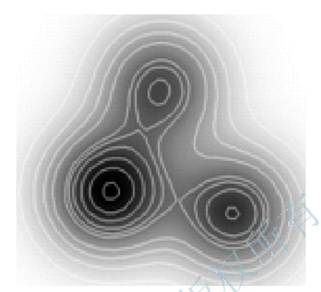
图 1.10　反射图与光照方向的关系(续)

(2) 等照度线也称等光亮度轮廓，是指图像中亮度为恒定常数的曲线，如图 1.11 所示。等照度线依赖于光源照明度和对象形状，若已知光源，可以根据等照度线图揭示出形状信息。人类视觉系统正是使用了这种线索。



(a) 原图

图 1.11　等照度线



(b) 等照度线图

图 1.11 等照度线(续)

2. SFS 理论基础

SFS 是典型的非适定问题，其数学原理如图 1.12 所示。

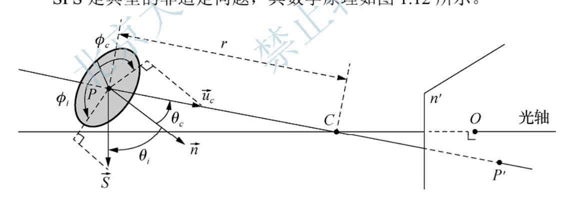


图 1.12 SFS 数学原理

SFS 的输入为图像明暗渐变性，输出为r、θ_c、ϕ_c、θ_i、ϕ_i。最著名的 SFS 模型是光程方程(也称短时距方程)，基于一阶偏微分方程和非线性方程理论。需要假定一些约束条件，如照片是清晰的、相机已光测度标定、前置光源是平行光且分布均匀、表面是 Lambertian 模型且可微等，且忽略内部交互反射现象。图 1.13 所示的图像均符合上

述假定条件。

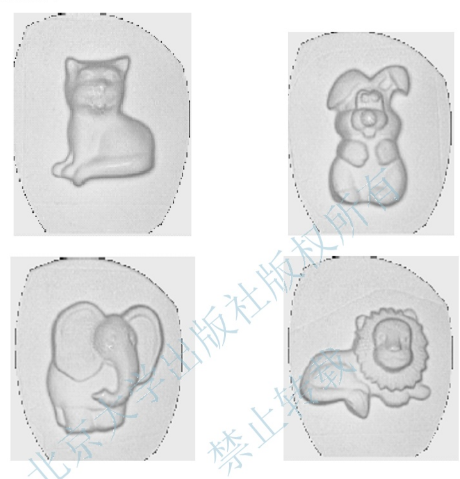


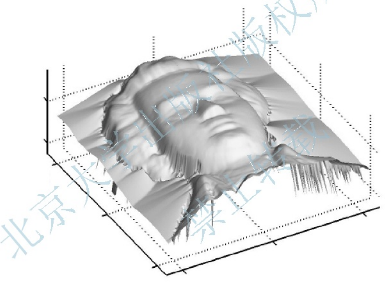
图 1.13　符合光程方程的图像示例

3. SFS 技术基础

Horn[38]首次系统阐述了 SFS 表面重构模式，即假设已知一个无穷远光源，物体表面呈 Lambertian 反射率和均匀分布的恒定反照率，平滑表面且不存在内部交互反射，并已知边界约束条件。在这样的假定下，图像中某像素的灰度级仅依赖于光源方向和表面法向。SFS 的基本任务是，基于给定的单幅图像，恢复表面每一点的光源方向和表面形状，如图 1.14 所示。



(a) 原图



(b) SFS结果

图 1.14　SFS 表面重构示例

SFS 方法实质上定义了表面形状与明暗之间的映射关系，即反射率函数 $I_{x,y} = R(p,q)$，其中 $I_{x,y}$ 表示图像像素强度，$p \doteq z_x$，$q \doteq z_y$，z 为深度，(x,y) 为 3D 对象空间坐标的投影。若假设为正投影，且表面为 Lambertian 反射模型，则得到下面的亮度约束。

$$I_{x,y} \propto \rho_{x,y} \frac{1 - pl - qk}{\sqrt{p^2 + q^2 + 1}\sqrt{l^2 + k^2 + 1}} \tag{1.2}$$

其中，$(l,k,1)/\|(l,k,1)\|$ 是光源方向。SFS 的基本任务是，估计式(1.2)中的未知量，即表面反照率 $\rho_{x,y}$ 及表面深度 $z(x,y)$。由于仅知道图像强度，因此，同时估计反照率和表面深度是一个病态问题。一个惯例是假设一个常量表面反照率，但仅根据这个单纯的假设来从实际图像中估计深度信息是非常脆弱的。因此，通常还需要另外强加一些约束条件。例如，Shimshoni[39]和 Chelloppa[40,41]等侧重研究对称物体的 SFS 重构问题，如人脸、车辆、建筑物等规则对象的重构。

但是真实图像并非总是 Lambertian 类型，即使假设其是 Lambertian 反射的，且已知光源方向，同时光亮度可以描述为表面形状和光源方向的函数，仍然难以轻易得到理想的结果。因为，假设用表面法向来表示表面形状，得到一个含有三个未知量的线性方程；而若采用表面梯度来表示表面形状，则得到一个含有两个未知量的非线性方程。所以，若想从 SFS 得到唯一解，必须附加额外的假设条件[42,43,44]。表 1-3 列出了较典型的三类方法的求解模式。

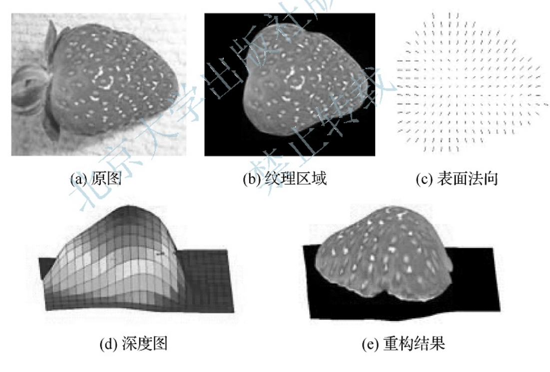
表 1-3　三类经典 SFS 方法的性能比较

方法	需要边界条件	解类型	执行难度
偏微分方程法	需要	精确解 $\varSigma: \boldsymbol{F}(\varSigma) = \varPhi$	难
最优化法	不需要	逼近解 $\varSigma: \boldsymbol{F}(\varSigma) \approx \varPhi$	难
模型逼近法	需要	逼近解 $\varSigma: \boldsymbol{F}(\varSigma) \approx \varPhi$	易

还有一种类似 SFS 的重构方法是 SFPS，其与 SFS 的差别仅在于输入图像的数目，SFPS 是根据同一场景的多幅图像重构，即视向不变，而光源方向不同。

1.3.2　SFT

SFT 模式在多年前由 Gibson [45]首次提出，指出纹理可以提供重要的形状线索。SFT 通常度量图像中的纹理畸变，重构 3D 表面坐标[46,47,48]。严格来讲，某点处的表面由表面法向与该点处的表面曲率确定。由于曲率方程相当复杂，不易于计算，因此，一般的 SFT 方法主要关注通过表面法向来重构表面形状，即根据纹理元素的畸变信息恢复表面法向。图 1.15 所示是一个葡萄的 SFT 结果。


(a) 原图　(b) 纹理区域　(c) 表面法向

(d) 深度图　(e) 重构结果

图 1.15　使用 SFT 恢复表面法向示例

SFT 的输出一般是表面法向的深度图。SFT 也是严重约束不足问题，要实现从图像的纹理畸变恢复表面法向，大多需要一些假设。例如，假设相机是已标定的或是正投影，也可假设纹理的类型，如各项

同性的、正则性的、同质的等。

所以，SFT 主要用于解决特定类型的纹理表面[49]。本书不进行全面探讨，在第 4 章仅对二进制歧义性纹理做进一步讨论。

1.4　歧义性问题

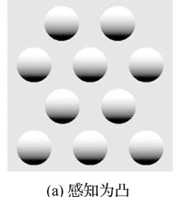
生理学研究表明，人类视觉总是试图将图像解释为 3D 表面结构[50,51,52,53,54,55]。人类的 3D 感知能力是人类视觉的核心，在对世界的认知与识别中扮演着重要角色。但是，由于单幅图像 3D 重构的固有歧义性，人们从不同视点观察得到的同一表面的估计结果也可能是不相同的。

计算机视觉的基础研究表明[56]，对于任意一个图像对，存在一个 Lambertian 表面，即在两种不同光照条件下产生两幅图像[57]。这意味着，不存在表面图像的任何一个非平凡函数，能在照度变化时保持完全不变。因此，单幅图像的 3D 重构问题是严重约束不足的，不可避免地存在歧义性，必须增加一些假设条件，如深度、明暗、纹理畸变等，这些性质在投影图像中留下了一些有用的暗示线索。基于单幅图像的 3D 重构方法正是充分利用了这些遗留痕迹。

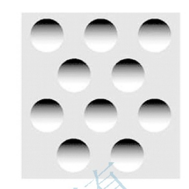
歧义性大致可分为局部歧义性和全局歧义性两类。在局部区域，关于深度的暗示信息一般具有很大的歧义性，因此，若想从一个较小的图像窗口中恢复 3D 形状是不可能的。

图 1.16 所示即为局部歧义性，仅通过将图案上下倒置，即可产生

不同的感知效果。



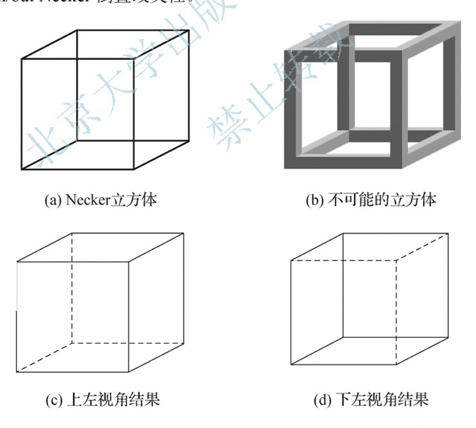
(a) 感知为凸



(b) 感知为凹

图 1.16　局部歧义性示例

全局歧义性的情况一般存在于整体性重构方案中。图 1.17 所示的是 in/out Necker 倒置歧义性。



(a) Necker立方体　　(b) 不可能的立方体

(c) 上左视角结果　　(d) 下左视角结果

图 1.17　全局歧义性示例——in/out Necker 倒置歧义性

Necker 立方体由 Necker 等于 1832 年首次提出[58]。由于 Necker 立方体具有很强的歧义性，常用于测试人类视觉系统的计算机模型，判断其能否达到和人类视觉系统感知一样的解释结果。如果立方体棱边以不相容方式交叉，则不可能构成几何物体，如图 1.17(b)所示。人类通常不会看到立方体的一个不相容的结果。对于如图 1.17(a)所示的 Necker 立方体，大多数人首先看到的是如图 1.17(c)所示的面，这可能因为人们是从上面俯视物体，而且是从左到右扫描物体。如果换一种观察角度，会得到如图 1.17(d)所示的另一种效果。

歧义性研究很重要，因为歧义性揭示了基于特定假设的自底向上重构方法的局限性，而不管所用的优化过程有多么成功。正是由于歧义性的广泛存在，使得人们对单图像 3D 重构的研究热情在很长一段时间内降到冰点，直到最近数年才又逐步得到重视。

1.5 基本约束模式

本书主要解决单幅图像的表面重构问题。由于上述歧义性的广泛存在，因此需要增加必需的约束条件。下面简单介绍表面重构的基本约束模式。

通常将一个表面表示为一个函数 $z(x,y)$，包含 N 个点 (x_i, y_i, z_i)，通过一个基函数 $\boldsymbol{B}$ 与一个参数向量 $\boldsymbol{v}$ 关联，即 $(z_i,\cdots,z_N)^{\mathrm{T}} = \boldsymbol{B}\boldsymbol{v}$。例如，在一个栅格上的一个高度图具有一个平凡基 $\boldsymbol{B} = \boldsymbol{I}$，$\boldsymbol{v} = (z_1,\cdots,z_N)^{\mathrm{T}}$。

基函数的选择主要考虑在基函数数量与保真度之间取一个折中。

在数学上，这种线性表示是相似的，它们之间可以相互转换。也就是说，如果

$$(z_1,\cdots,z_N)^{\mathrm{T}} \approx \boldsymbol{B}_1\boldsymbol{v}_1 \approx \boldsymbol{B}_2\boldsymbol{v}_2 \tag{1.3}$$

利用伪逆算子得到

$$\boldsymbol{v}_1 \approx \boldsymbol{B}_1^+\boldsymbol{B}_2\boldsymbol{v}_2 \tag{1.4}$$

这样，在一个线性表示中指定的约束，可投射到另一个线性表示。线性表示的好处是，表面的深度和导数都是简单的线性函数。例如，在一个栅格上，可以用有限差 $\boldsymbol{p}_{i,j} = \boldsymbol{z}_{i+1,j} - \boldsymbol{z}_{i,j}$，$\boldsymbol{q}_{i,j} = \boldsymbol{z}_{i,j+1} - \boldsymbol{z}_{i,j}$ 来逼近表面导数。

表面的法向是一个向量，方向为$(-p,-q,1)$。有时候，也可用 slant 角 σ 和 tilt 角 τ 来表示单位法向向量，如图 1.18 所示。

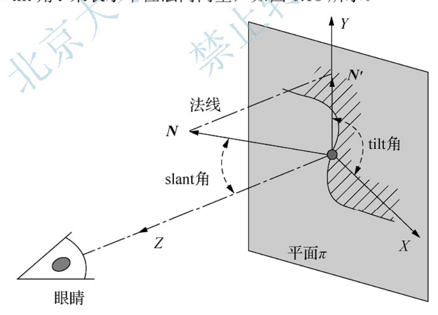


图 1.18　slant 角和 tilt 角

slant 角为视线与法向 $\boldsymbol{N}$ 之间的夹角 σ。平面 π 与视线垂直，设交点即为坐标系的原点，法向 $\boldsymbol{N}$ 在平面上的投影为 $\boldsymbol{N}'$，$\boldsymbol{N}'$ 与 X 轴的夹角就是 tilt 角 τ。这种 (σ,τ) 表示类似于地理学上的(维度,经度)。在这样的表示中，单位法向向量为 $(\sin\sigma\cos\tau,\sin\sigma\sin\tau,\cos\sigma)$。

1. 线性约束

线性约束是迄今研究的最多的一类约束，如表 1-4 所示。

表 1-4 常见的线性约束

约束	公式
深度	$z_i = z_0$
共面点	$z_i = ax_i + by_i + d$
法向	$p_i = p_0, q_i = q_0$
法向摆角	$q_i = p_i \tan\tau$

2. 离散约束

离散约束最典型的是二进制(符号)的歧义性，如图 1.17 中所示的 Necker in/out 倒置歧义性。这种歧义性一般出现在 SFT 分析中。另一种重要的二进制歧义性是 Perkins 立方体角[59,60]，即三个互垂线的交点，如图 1.19 所示。

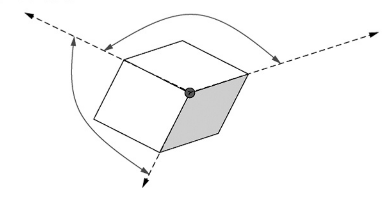


图 1.19 Perkins 的立方体角

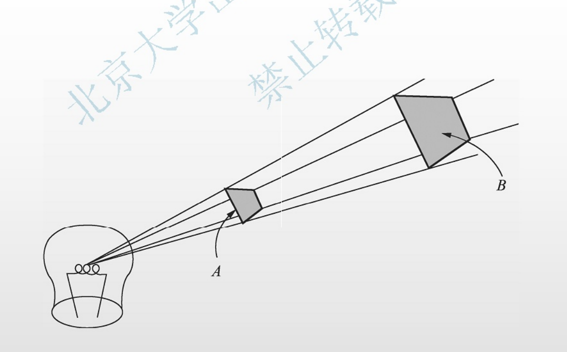
3. 二次约束

例如，法向分别为$(-p_1,-q_1,1)$和$(-p_2,-q_2,1)$的两个平坦削面的正交约束，可以写成$p_1q_1+p_2q_2+1=0$。这部分内容将在第 5 章详细介绍。





第 2 章 3D 表面成像及优化方法



日常所见的图像大多数是光学成像的结果，如太阳光、白炽光等。根据图 1.2，要获得一个 3D 场景所对应的 2D 图像，要经过复杂的几何变换、射影变换等过程，并且要考虑具体所用的相机模型。在第 3、4、5 章介绍的基于单幅图像的 3D 重构技术中，将会使用到一些特殊的数学分析方法，如半定规划、同伦分析法等。为便于后续章节的叙述与阅读，本章将简要介绍这些基础知识，重点阐述这些知识与本书拟解决问题之间的内在联系。

2.1 3D 表面概念

图 2.1 描述了 3D 表面的直观概念。一个表面可表示为方程 $z = f(x, y)$，包含了表面方向和深度信息。

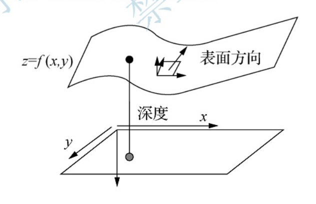


图 2.1 一个 3D 表面

表面的法向量表示为

$$\boldsymbol{n} = \frac{\partial z}{\partial x} \cdot \frac{\partial z}{\partial y} \tag{2.1}$$

2.1.1 表面的形式化定义

下面给出 3D 表面的形式化定义。

定义 2.1(3D 表面)

三维空间 $\mathbb{R}^3$ 中的一个平滑表面定义为一个子集 $X \subset \mathbb{R}^3$，其中的每一个点对应一个邻域 $U \subset X$。存在一个映射 $\boldsymbol{r}: V \to \mathbb{R}^3$ 满足下述条件，其中 $V \subseteq \mathbb{R}^2$。

(1) $\boldsymbol{r}: V \to U$ 是一个拓扑同构的。

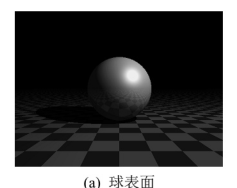
(2) $\boldsymbol{r}(u,v) = (x(u,v, y(u,v), z(u,v)))$ 具有全阶导数。

(3) 每一点的偏导数 $\boldsymbol{r}_u = \partial \boldsymbol{r} / \partial u, \boldsymbol{r}_v = \partial \boldsymbol{r} / \partial v$ 是线性无关的。

定义 2.1 从拓扑层面描述了一个表面，$\boldsymbol{r}$ 定义了一个同质的映射 $\varphi_U: U \to V$。上述条件(2)和(3)可用隐函数定理来理解，即若将 $\mathbb{R}^3$ 中的变量局部倒转，将直接产生 3D 表面。对于任意两个开放集 U, U'，得到一个从一个开放集到另一个开放集的平滑倒转映射。这意味着每个映射 $\varphi_U, \varphi_U^{-1}$ 是一个平滑的拓扑同构。这样就可得到下面的 3D 表面的抽象化定义。

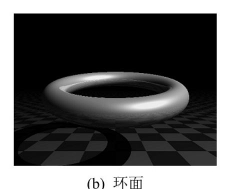
定义 2.2(表面的抽象定义)

若一个表面拥有一类拓扑同构 φ_U，其中每个映射 $\varphi_U, \varphi_U^{-1}$ 是一个平滑可逆的拓扑同构，则称为平滑表面。图 2.2 所示为较常见的平滑表面。



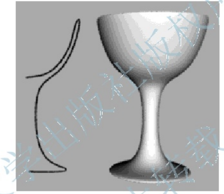
(a) 球表面

$$\boldsymbol{r}(u,v) = a\sin u \sin v_i + a\cos u \sin v_j + a\cos v_k$$



(b) 环面

$$\boldsymbol{r}(u,v) = (a + b\cos u)(\cos v_i + \sin v_j) + b\cos u_k$$



(c) 旋转表面

$$\boldsymbol{r}(u,v) = f(u)(\cos v_i + \sin v_j)$$

图 2.2 常见平滑表面

2.1.2 第一基本型

定义 2.3(表面上的平滑曲线)

一条位于表面上的平滑曲线是一个映射，即下面的映射

$$t \mapsto (u(t), v(t))$$

且具有全阶导数，则 $\gamma(t) = \boldsymbol{r}(u(t), v(t))$ 称为 $\mathbb{R}^3$ 上的一条平滑曲线，如图 2.3 所示。

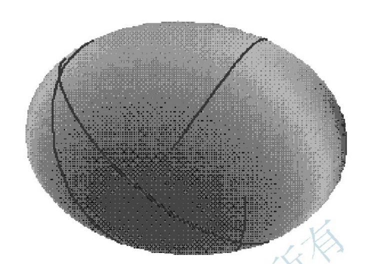


图 2.3　位于表面的平滑曲线

一条参数化的曲线意味着 $u(t), v(t)$ 有全阶导数，且 $\gamma' = \boldsymbol{r}_u u' + \boldsymbol{r}_v v' \neq 0$。根据上述定义，$\boldsymbol{r}_u, \boldsymbol{r}_v$ 是线性无关的，因此，该条件等价于 $(u', v') \neq 0$。

在一条曲线上，从 a 点到 b 点的弧长定义为

$$\begin{aligned}\int_a^b |\gamma'(t)| \mathrm{d}t &= \int_a^b \sqrt{\gamma' \cdot \gamma'} \mathrm{d}t = \int_a^b \sqrt{(\boldsymbol{r}_u u' + \boldsymbol{r}_v v')(\boldsymbol{r}_u u' + \boldsymbol{r}_v v')} \mathrm{d}t \\ &= \int_a^b \sqrt{Eu'2Fu'v' + Gv'} \mathrm{d}t\end{aligned} \tag{2.2}$$

定义 2.4 (第一基本型)

在 $\mathbb{R}^3$ 域中，一个表面的第一基本型表示为

$$E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + G\mathrm{d}v^2 \tag{2.3}$$

其中，$E = \boldsymbol{r}_u \cdot \boldsymbol{r}_u$，$F = \boldsymbol{r}_u \cdot \boldsymbol{r}_v$，$G = \boldsymbol{r}_v \cdot \boldsymbol{r}_v$。第一基本型恰好是二次项形式 $Q(\boldsymbol{v}, \boldsymbol{v}) = \boldsymbol{v} \cdot \boldsymbol{v}$，对应一个切空间，$\boldsymbol{r}_u, \boldsymbol{r}_v$ 为该空间的基。第一基本型描述了表面的内在几何特性。

2.1.3 第二基本型

将表面 $\boldsymbol{r}(u,v)$ 沿着其法向移动一个 t 距离，从而得到一个表面簇

$$\boldsymbol{R}(u,v,t)=\boldsymbol{r}(u,v)-t\boldsymbol{n}(u,v) \tag{2.4}$$

其中，$\boldsymbol{R}_u=\boldsymbol{r}_u-t\boldsymbol{n}_u$，$\boldsymbol{R}_v=\boldsymbol{r}_v-t\boldsymbol{n}_v$。

若已知第一基本型 $E\mathrm{d}u^2+2F\mathrm{d}u\mathrm{d}v+G\mathrm{d}v^2$，则可按式（2.5）计算

$$\frac{1}{2}\frac{\partial}{\partial t}(E\mathrm{d}u^2+2F\mathrm{d}u\mathrm{d}v+G\mathrm{d}v^2)\Big|_{t=0}=-(\boldsymbol{r}_u\cdot\boldsymbol{n}_u\mathrm{d}u^2+(\boldsymbol{r}_u\cdot\boldsymbol{n}_v+\boldsymbol{r}_v\cdot\boldsymbol{n}_u)\mathrm{d}u\mathrm{d}v \\ +\boldsymbol{r}_v\cdot\boldsymbol{n}_v\mathrm{d}v^2) \tag{2.5}$$

式（2.5）的右端一般称为第二基本型。

定义 2.5(第二基本型)

$\mathbb{R}^3$ 中一个表面的第二基本型表示为

$$L\mathrm{d}u^2+2M\mathrm{d}u\mathrm{d}v+N\mathrm{d}v^2$$

其中，$L=\boldsymbol{r}_{uu}\cdot\boldsymbol{n}$，$M=\boldsymbol{r}_{uv}\cdot\boldsymbol{n}$，$N=\boldsymbol{r}_{vv}\cdot\boldsymbol{n}$。

定理 2.1(表面是平面的一部分)

如果一个表面的第二基本型消失了，则该表面就是一个平面的一部分，称为局部平面。

证明：

若表面的第二基本型消失，即 $\boldsymbol{r}_u\cdot\boldsymbol{n}_u=0=\boldsymbol{r}_v\cdot\boldsymbol{n}_u=\boldsymbol{r}_u\cdot\boldsymbol{n}_v=\boldsymbol{r}_v\cdot\boldsymbol{n}_v$，则 $\boldsymbol{n}_u=\boldsymbol{n}_v=0$。

由于 $\boldsymbol{n}_u$、$\boldsymbol{n}_v$ 垂直于 $\boldsymbol{n}$，因此是 $\boldsymbol{r}_u,\boldsymbol{r}_v$ 的线性组合。这意味着 $\boldsymbol{n}$ 是常

量，即

$$\begin{cases}(\boldsymbol{r}\cdot\boldsymbol{n})_u=\boldsymbol{r}_u\cdot\boldsymbol{n}=0\\(\boldsymbol{r}\cdot\boldsymbol{n})_v=\boldsymbol{r}_v\cdot\boldsymbol{n}=0\end{cases}$$

因此，$\boldsymbol{r}\cdot\boldsymbol{n}$ 是常量，即为平面方程。

□

定理 2.2(闭合表面上的点的正定性)

对于 $\mathbb{R}^3$ 中的任意一个闭合表面 X，含有一些点，在这些点上的第二基本型是正定的。

证明：

由于 X 是闭合的，因此为有界，可被一个大球围绕。可不断地压瘪该球，直到半径 R 边缘停靠在 X 的一个点上。若将 X 描述为 f 函数图，则有

$$f-(R-\sqrt{R^2-x^2-y^2})\geqslant 0 \tag{2.6}$$

式（2.6）泰勒级数的第一个非零项为

$$\frac{1}{2}(f_{xx}x^2+2f_{xy}xy+f_{yy}y^2)-\frac{1}{2R}(x^2+y^2) \tag{2.7}$$

因此，有 $Lx^2+2Mxy+Ny^2\geqslant\frac{1}{R}(x^2+y^2)>0$，即为正定的。

□

2.2 光照模型

2.2.1 光源与辐照度

假设一束光射线位于一个管带内，与管带的两端相交，但与侧边不相交，如图 2.4 所示。

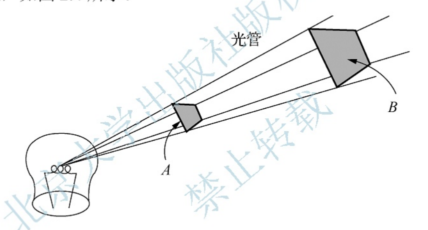


图 2.4 光照模型

1. 辐照度

辐照度用于度量光射线碰撞测试对象时所产生的能量 $I(\lambda,\vec{\boldsymbol{x}}_p,\vec{\boldsymbol{n}}_p)$，其中 $\vec{\boldsymbol{x}}_p$ 为碰撞点位置，$\vec{\boldsymbol{n}}_p$ 为表面法向，如图 2.5(a)所示。

2. 光源辐照度

设光源方向为 $\vec{\boldsymbol{L}}$，一般指从测试点目视光源的方向，通常定义为该方向上的单位向量。

如图 2.5(b)所示，光源方向等于表面法向，即 $\vec{\boldsymbol{n}}_p=\vec{\boldsymbol{L}}$，此时

$I(\lambda, \vec{x}_p, \vec{n}_p)$ 替换为 $I(\lambda, \vec{L})$ 。

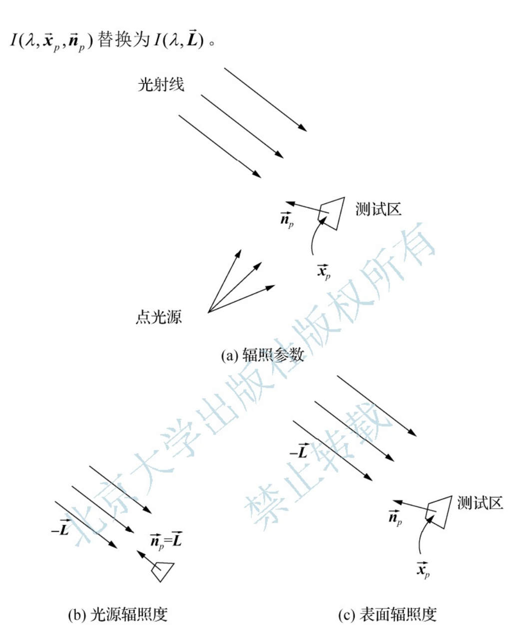


图 2.5　辐照过程

3. 表面辐照度

考虑 $\vec{n}_p \neq \vec{L}$ 的一般情况，如图 2.5(c)所示。光射线碰撞测试区的能量就是穿过测试区投影面积的能量，而投影面积与光线角度成比例，如图 2.6 所示。

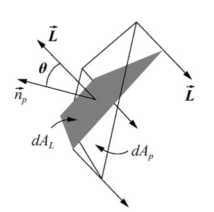


图 2.6　测试区的投影面积

投影面积为 $dA_L = |\cos(\theta)| dA_p = |\vec{n}_p \cdot \vec{L}| dA_p$。

如果是不透明表面，要求 $\vec{n}_p \cdot \vec{L} \geqslant 0$，否则，就是阴影。因此，表面辐照度为

$$I(\lambda, \vec{n}_p) = \lfloor \vec{n}_p \cdot \vec{L} \rfloor I(\lambda, \vec{L}) \tag{2.8}$$

4. 点光源下的表面辐照度

如图 2.7 所示，点光源的辐射强度 $R(\lambda, \vec{\theta}_p)$ 是指在 $\vec{\theta}_p$ 方向中碰撞到测试区的单位波长 λ、单位立体角的能量。

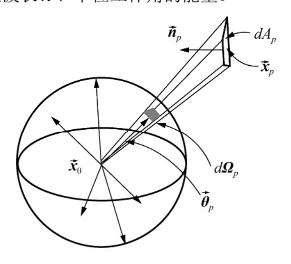


图 2.7　点光源辐照度

所谓立体角，定义为测试区在单位球上的投影，即

$$d\boldsymbol{\Omega}_p = \frac{\left|\vec{\boldsymbol{n}}_p \cdot \vec{\boldsymbol{L}}\right|}{\left\|\vec{\boldsymbol{x}}_p - \vec{\boldsymbol{x}}_0\right\|^2} dA_p$$

其中，$\vec{\boldsymbol{L}} = (\vec{\boldsymbol{x}}_0 - \vec{\boldsymbol{x}}_p) / \left\|\vec{\boldsymbol{x}}_0 - \vec{\boldsymbol{x}}_p\right\|$。

对于测试区，其表面辐照度为

$$I(\lambda, \vec{\boldsymbol{x}}_p, \vec{\boldsymbol{n}}_p) = \frac{\left[R(\lambda, \vec{\boldsymbol{\theta}}_p) d\boldsymbol{\Omega}_p\right]}{dA_p} = \frac{\left|\vec{\boldsymbol{n}}_p \cdot \vec{\boldsymbol{L}}\right|}{\left\|\vec{\boldsymbol{x}}_p - \vec{\boldsymbol{x}}_0\right\|^2} R(\lambda, \vec{\boldsymbol{\theta}}_p) \qquad (2.9)$$

5. 分布式光源辐照度

点光源情况是一种理想假设，实际情况下多为有限个光源的组合，如图 2.8 所示。

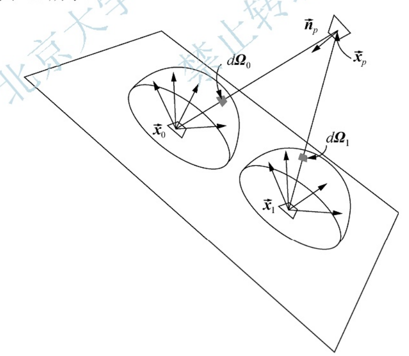


图 2.8　分布式光源

其中一个测试区的辐照度为

$$\frac{\lfloor \vec{\boldsymbol{n}}_p \cdot \vec{\boldsymbol{L}}(\vec{\boldsymbol{x}},\vec{\boldsymbol{x}}_p) \rfloor}{\|\vec{\boldsymbol{x}}-\vec{\boldsymbol{x}}_p\|^2} R(\lambda,\vec{\boldsymbol{x}},\vec{\boldsymbol{\theta}}(\vec{\boldsymbol{x}}_p,\vec{\boldsymbol{x}}))d\vec{\boldsymbol{x}} \tag{2.10}$$

一般情况下，这些光源之间是一种线性组合，可以进行叠加运算，结果相当于单光源。因此，后面的章节都是基于单光源展开讨论，这种假设是合理的。

2.2.2 表面反射模型

表面反射模型如图2.9所示。

图2.9 表面反射模型

表面某点的图像强度为$I_o = f(\vec{\boldsymbol{N}},k_d,I_i)$，其中$\vec{\boldsymbol{N}}$为表面法向，$k_d$为表面反射率，$I_i$为入射光强度。

1. *漫反射*

设一个无穷处光源，方向为$\vec{\boldsymbol{L}}$，表面辐照度是$I(\lambda,\vec{\boldsymbol{N}})$，其中$\vec{\boldsymbol{N}}$是

表面法向，$\vec{V}$ 为视向，如图 2.10 所示。

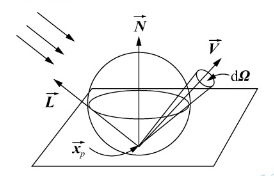


图 2.10　漫反射模型

单位面积表面的反射强度为

$$R(\lambda,\vec{x}_p,\vec{V};\vec{N})=r(\lambda,\vec{L},\vec{V})\left\lfloor\vec{N}\cdot\vec{V}\right\rfloor I(\lambda,\vec{N})$$

其中，$r(\lambda,\vec{L},\vec{V})$ 称为双向反射分布函数(Bidirectional Reflectance Distribution Function，BRDF)。

若假设 BRDF 对于 $\vec{L},\vec{V}$ 保持恒定，则 $r(\lambda,\vec{L},\vec{V})=r(\lambda)$，称为 Lambertian 逼近模型。

2. 镜面反射

考虑一个无穷远处的光源，其辐照度为 $I^i(\lambda,\vec{L})$，其中的上标 i 表示入射辐照度。

图 2.11 中的 $\vec{M}$ 是镜面反射方向，可表示为 $\vec{M}=-\vec{L}+2\vec{N}\left[\vec{N}\cdot\vec{L}\right]$，此时的反射辐照度为

$$I^r(\lambda,\vec{M})=F(\lambda,\vec{L})I^i(\lambda,\vec{L}) \tag{2.11}$$

其中的 $F(\lambda,\vec{L})$ 是 Fresnel 项，指明光谱分布的变化。

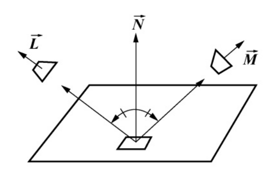


图 2.11　镜面反射模型

2.2.3　Lambertian 模型

由上述可知，表面辐照度与光源辐照度的关系为

$$I(\lambda,\vec{N})=\lfloor\vec{N}\cdot\vec{L}\rfloor I(\lambda,\vec{L})$$

横截面区域在$\vec{V}$上的垂直投影为dA_V，如图 2.12 所示。

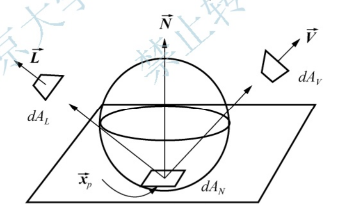
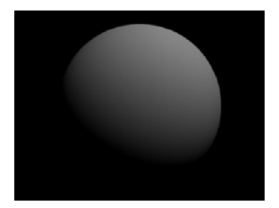


图 2.12　Lambertian 模型

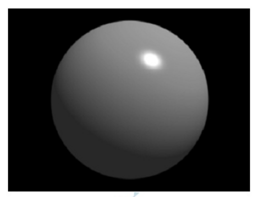
与反射表面的对应区域为$dA_V=\left|\vec{N}\cdot\vec{V}\right|dA_N$，则有

$$R(\lambda,\vec{x}_p,\vec{V})=r(\lambda)\lfloor\vec{N}\cdot\vec{L}\rfloor I(\lambda,\vec{L}) \tag{2.12}$$

这意味着其不依赖于视向$\vec{V}$，如图 2.13 所示。



(a)　Lambertian 模型



(b) Lambertian 模型+镜面

图 2.13　球面 Lambertian 模型

2.2.4　梯度空间

如图 2.14 所示，表面法向表示为平面上的一点(p,q)，而光源方向表示为另一个点(p_s,q_s)。表面重构的目标之一就是根据图像光亮度$I(x,y)$，估计梯度空间(p,q)。

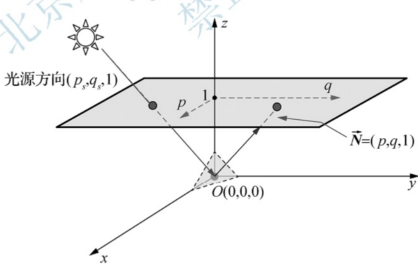


图 2.14　梯度空间示意图

单位法向量为

$$\vec{n}=\frac{\vec{N}}{\|\vec{N}\|}=\frac{(p,q,1)}{\sqrt{p^2+q^2+1}}$$

光源法向为

$$\vec{s}=\frac{\vec{N}_s}{\|\vec{N}_s\|}=\frac{(p_s,q_s,1)}{\sqrt{p_s^2+q_s^2+1}}$$

它们的夹角为

$$\cos\vartheta_i=\vec{n}\cdot\vec{s}=\frac{1+pp_s+qq_s}{\sqrt{1+p^2+q^2}\sqrt{1+p_s^2+q_s^2}}$$

2.3 光学成像原理

要实现基于图像的 3D 重构，首先必须深入了解成像设备的物理性能，如相机的内外参数等。因此，下面介绍最常用的针孔相机的成像模型。

2.3.1 相机模型

针孔相机可实现将一个 3D 场景及其所包含的物理对象映射到一个像平面，从而获得一幅 2D 图像，这个过程称为投影，如图 2.15 所示。

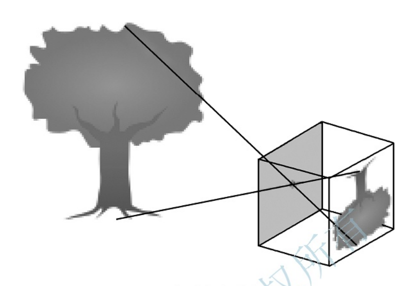


图 2.15　针孔相机的投影成像

投影表示为一个 3×4 的矩阵 $\boldsymbol{F}$，描述了 $\mathbb{P}^3$ 空间中的场景和 $\mathbb{P}^2$ 空间中图像之间的关系：$m \propto \boldsymbol{F}M$，其中

$$\boldsymbol{F}=\begin{bmatrix} f_{11} & f_{12} & f_{13} & f_{14} \\ f_{21} & f_{22} & f_{23} & f_{24} \\ f_{31} & f_{32} & m_{33} & f_{34} \end{bmatrix} \tag{2.13}$$

其中，P, p 分别为 3D 点及其投影的 2D 点，均采用齐次坐标；矩阵元素 $f_{ij} \in \mathbb{R}$ 。

图 1.2 所示的相机成像模型描述了场景的透视投影，使用若干参数，这些参数表示相机的内部参数以及与场景的关系。投影中心为 O_C；3D 空间点 M 在相机坐标系中 $< XYZ, O_C >$ 的坐标为 (X_c, Y_c, Z_c)；主轴与像平面的交点 O_u 称为主点，一般是像平面坐标系 $< xy, o_u >$ 的原点。

M 投影为图像点 m，表示为

$$x = f\frac{X_c}{Z_c}\ ,\ y = f\frac{Y_c}{Z_c} \tag{2.14}$$

其中，f 为投影中心到像平面的距离，称为焦距。

需要注意以下几点。

(1) 对于数字照相机，如 CCD 或 CMOS 相机，像平面是像素阵列，单个像素的水平和垂直长度可能并不相同，其比值称为纵横比。

(2) 同样，由于制造工艺问题，像素也许并不一定是矩形，也可能是平行四边形或梯形[61]，即存在倾斜，倾斜因子一般不为零，但是近似为零，如图 2.16 所示。

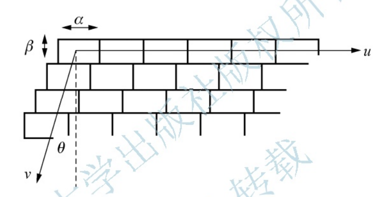


图 2.16　倾斜因子和纵横比

(3) 一般情况下，主点并不位于像平面的原点，如图 2.17 所示。

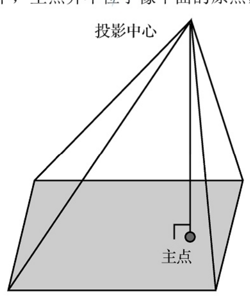


图 2.17　主点不位于像平面原点

因此，像平面上得到的一个像素点 $\tilde{p}(u,v)$ 可表示为

$$\begin{cases} u = \dfrac{x}{\alpha} + \dfrac{y}{\beta}\tan\theta + u_o \\ v = \dfrac{y}{\beta} + v_o \end{cases} \tag{2.15}$$

其中，(u_o, v_o) 是主点在坐标系 $< xy, o_u >$ 中的位置坐标。α 为像素的水平长度，β 为像素的垂直长度。θ 为 u,v 方向的夹角。

如果场景和图像点表示为齐次坐标，那么，中心投影可简单表示为线性映射。组合式(2.14)和式(2.15)得到

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \simeq \begin{bmatrix} f\dfrac{\beta}{\alpha} & f\dfrac{\beta}{\alpha}\tan\theta & u_o & 0 \\ & f & v_o & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{2.16}$$

1. 内参数

式(2.16)可以改写为

$$p \simeq \boldsymbol{K}[I \mid O]P_c \tag{2.17}$$

其中，

$$\boldsymbol{K} = \begin{bmatrix} f\dfrac{\beta}{\alpha} & f\dfrac{\beta}{\alpha}\tan\theta & u_o \\ & f & v_o \\ & & 1 \end{bmatrix} = \begin{bmatrix} f_u & s & u_o \\ & f_v & v_o \\ & & 1 \end{bmatrix} \tag{2.18}$$

式(2.18)中的 f_u, f_v 分别表示焦距在像平面坐标轴 u,v 上的投影焦距，纵横比为 f_u / f_v， s 为倾斜因子。一般地， $\boldsymbol{K}$ 称为相机的内参数。

2. 外参数

上面假设场景中的 3D 点 $\tilde{P}_c(X_c, Y_c, Z_c)$ 是用相机坐标系来描述，事实上，一般描述为世界坐标系 $<XYZ,O>$。这种情况下，为了计算的有效性，所有数据点包括相机中心都必须定义在世界坐标系中。

假设相机的光学中心为 $\tilde{C}$，则有

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \boldsymbol{R}\left(\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} - \tilde{C}\right) \tag{2.19}$$

其中，$\boldsymbol{R}$ 是旋转矩阵，表示了相机坐标系对于世界坐标系的方位。

式(2.19)意味着，如果世界坐标系中的 3D 点和 $\boldsymbol{R}, \tilde{C}$ 是有效的，则该点可定义在相机坐标系。参数 $\{\boldsymbol{R}, \tilde{C}\}$ 就称为相机的外参数。相机的外参数如图 2.18 所示。

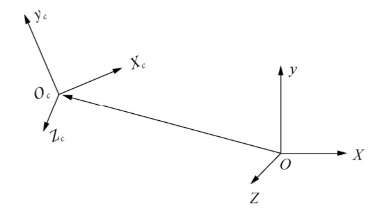


图 2.18　相机的外参数

综上所述，得到

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boldsymbol{KR}[I \mid -\tilde{C}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \boldsymbol{M} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.20}$$

其中，

$$\boldsymbol{M} = \boldsymbol{KR}[I \mid -\tilde{C}] \tag{2.21}$$

式(2.21)称为相机的投影矩阵，也称基础矩阵。

如果知道投影矩阵 $\boldsymbol{M}$，就可能由此分解得到相机的内参数 $\boldsymbol{K}$ 和外参数 $\{\boldsymbol{R}, \tilde{C}\}$。

2.3.2　投影变换

一个变换可以代数化定义为一个矩阵。

1. 射影变换

射影变换是一个线性变换，一般地会改变几何形状的角度、距离或比率。能确保恒定不变的是奇异性、关联性、交比。从 $\mathbb{P}^n$ 到 $\mathbb{P}^m$ 的射影变换，映射每个点 $P \in \mathbb{P}^n$ 到 $\mathbb{P}^m$ 中的一个新的位置：$P \mapsto \boldsymbol{T}P$，其中 $\boldsymbol{T}$ 是一个 $(m+1)\times(n+1)$ 的矩阵。

(1) 单应性。单应性是指在同一空间内的一个射影变换 $\boldsymbol{H}: \mathbb{P}^n \to \mathbb{P}^m$，即 $P \mapsto \boldsymbol{H}P$。

这里的 $\boldsymbol{H}$ 是一个 $(n+1)\times(n+1)$ 的非奇异矩阵。

在 P^3 空间，$\boldsymbol{H}$ 为 4×4 的矩阵，具有 15 个自由变量。

$$\boldsymbol{H} \simeq \begin{bmatrix} \boldsymbol{A} & \boldsymbol{t} \\ \boldsymbol{v}^{\mathrm{T}} & 1 \end{bmatrix} \tag{2.22}$$

其中，$A_{3\times3}$ 是非奇异矩阵；$\boldsymbol{t},\boldsymbol{v}$ 为 3-向量。

(2) 对偶空间的单应性。

当 P^n 空间的某个点经历 $\boldsymbol{H}$ 变换后，得到的对偶图元 $L\in\mathrm{P}^{*n}$，对于 2D 和 3D 空间，则分别表现为直线和平面。

对偶变换为 $\boldsymbol{H}^*:\mathrm{P}^{*n}\to\mathrm{P}^{*n}$，即 $L\mapsto\boldsymbol{H}^{-\mathrm{T}}L$。

2. 仿射变换

仿射变换是一种特殊的射影变换，保留无穷远处全局恒定不变的超平面。因此，在 P^3 空间，一个仿射变换及其逆变换可表示为

$$\boldsymbol{T}_A \simeq \begin{bmatrix} \boldsymbol{A} & \boldsymbol{t} \\ 0^{\mathrm{T}} & 1 \end{bmatrix}, \boldsymbol{T}_A^{-1} \simeq \begin{bmatrix} \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{t} \\ 0^{\mathrm{T}} & 1 \end{bmatrix} \tag{2.23}$$

$\boldsymbol{T}_A$ 含有 12 个变量。

3. 相似性变换

相似性变换是一种特殊的放射变换，包含平移、旋转和缩放。这种变换保持任何方向的夹角、比率不变。在 P^3 空间，相似变换可表示为

$$\boldsymbol{T}_M \simeq \begin{bmatrix} \lambda\boldsymbol{R} & \boldsymbol{t} \\ 0^{\mathrm{T}} & 1 \end{bmatrix} \tag{2.24}$$

其中，$\boldsymbol{R}_{3\times3}$ 是旋转矩阵，$\boldsymbol{t}=(t_1,t_2,t_3)^{\mathrm{T}}$ 是一个平移向量。λ 是一个各向同性的非零比例因子。$\boldsymbol{T}_M$ 含有 7 个变量，即围绕三个轴的旋转、在三个方向上的平移、一个全局性的比例因子。

4. 欧氏变换

欧氏(或等量)变换是相似性变换的特殊情况，包含一个平移和一个旋转。欧氏变换为一个物体的刚体变换建立模型。这种变换保持任何方向的夹角和距离不变。

在 P^3 空间，欧氏变换可表示为

$$\boldsymbol{T}_E \simeq \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ 0^{\mathrm{T}} & 1 \end{bmatrix} \tag{2.25}$$

其中，$\boldsymbol{R}_{3\times3}$ 是一个旋转矩阵，且 $\det(\boldsymbol{R})=1$，所以 $\boldsymbol{R}^{-1}=\boldsymbol{R}^{\mathrm{T}}$。

$\boldsymbol{T}_E$ 含有 6 个变量，即围绕三个轴的旋转、在三个方向上的平移。

上述四种变换的总结参见表 2-1。

比较表 2-1 中的射影变换和仿射变换的矩阵，可以发现，它们的差异在于向量 $\boldsymbol{v}$，在仿射变换中，$\boldsymbol{v}$ 为零向量。

表 2-1　四种 3D 几何变换

变换种类	矩阵	畸变	不变量	变量数
射影变换	$\boldsymbol{T} \simeq \begin{bmatrix} \boldsymbol{A} & \boldsymbol{t} \\ \boldsymbol{v}^{\mathrm{T}} & 1 \end{bmatrix}$		交叉性 笔直性	15

续表

变换种类	矩阵	畸变	不变量	变量数
仿射变换	$\boldsymbol{T}_A \simeq \begin{bmatrix} \boldsymbol{A} & \boldsymbol{t} \\ 0^{\mathrm{T}} & 1 \end{bmatrix}$		平面平行性 体积的比率、质心 无穷远平面 Π_∞	12
相似变换	$\boldsymbol{T}_M \simeq \begin{bmatrix} \lambda\boldsymbol{R} & \boldsymbol{t} \\ 0^{\mathrm{T}} & 1 \end{bmatrix}$		任何方向的夹角、比率	7
等量变换	$\boldsymbol{T}_E \simeq \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ 0^{\mathrm{T}} & 1 \end{bmatrix}$		任何方向的夹角、距离	6

因此，放射变换保留了无穷远处的理想点 $(x,y,z,0)^{\mathrm{T}}$ 。

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{t} \\ 0^{\mathrm{T}} & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 0 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \\ 0 \end{bmatrix} \tag{2.26}$$

在射影变换下，理想点 $(x,y,z,0)^{\mathrm{T}}$ 映射到一个有限远点，看作为一个消隐点。

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{t} \\ \boldsymbol{v}^{\mathrm{T}} & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 0 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \\ \boldsymbol{v}_1 x + \boldsymbol{v}_2 y + \boldsymbol{v}_3 z \end{bmatrix} \tag{2.27}$$

2.4　同伦分析法

基于单幅图像的 3D 表面重构一般是 NP-难题，难以构成线性系统。而非线性系统远比线性系统的求解要困难得多，严重依赖于所采取的优化逼近方法。考察一个逼近分析方法对于解非线性系统求解结果的满意度，通常有以下两个标准。

(1) 是否总能提供有效的逼近分析。

(2) 能否确保逼近解对所有的物理参数均能保持较高的精确度。

典型的扰动技术[62,63,64,65,66,67]被广泛应用于科学和工程领域。扰动技术主要基于给定方程或初始、边界条件等中的物理参数，一般称为微扰量。一般地，扰动逼近表示为一系列微扰量，一个非线性问题可分解为无限多的线性子问题，而这些线性子问题完全可由给定的初始方程确定，尤其取决于微扰量出现的位置。扰动法比较简单，特别是当基于较小的物理参数，扰动逼近具有明确的物理含义。不过，并非每个非线性问题都含有这样的微扰量。而且，即使存在这样的物理参数，线性子问题也可能无解，或者因为过于复杂而仅有少量子问题可解。这样，扰动法就不能确保对任何非线性问题总能得到扰动逼近。更重要的是，扰动逼近仅对小的物理参数有效，这样，就不能保证一个扰动结果对所有的参数域都有效。这样看来，扰动法不满足上述两条考察标准。

为了克服扰动法的局限性，出现了一些经典的非扰动方法，如人造较小参数法[68]、δ-扩展法[69,70]、Adomian 分解法[71,72,73,74,75,76]等。

根本上讲，这些方法都是基于一个人造参数，将逼近解扩展为一系列这样的人造参数。人造小参数一般以这样的方式来使用，即对任何给定的非线性方程，总能获得逼近解。与扰动技术比较，这的确是一个较大的进步。但是，从理论上讲，存在不同的方式来设置这个人造参数，却没有任何理论来阐明究竟如何设置人造参数才能获得更好的逼近解。所以，这些非扰动法仅满足上述第(1)条，而不满足第(2)条。

2.4.1 早期 HAM

1992 年廖世俊[76]首次将拓扑学中的同伦概念[77]用于非线性微分方程的逼近分析，从而诞生了一种全新的优化逼近分析方法，即同伦分析法(Homotopy Analysis Technique，HAM)。

设一个给定的非线性微分方程为

$$N[u(x)] = 0, \quad x \in \Omega$$

其中，N 是一个非线性算子；$u(x)$ 是一个未知函数。廖世俊据此构造了一个仅含一个参数的方程簇，嵌入参数 $q \in [0,1]$，称为零阶形变方程。

$$(1-q)L[\phi(x;q) - u_0(x)] + qN[\phi(x;q)] = 0, \quad x \in \Omega, q \in [0,1] \qquad (2.28)$$

其中，L 是一个辅助线性算子；$u_0(x)$ 是一个初始值。

理论上，拓扑学中的同伦概念为选择 L 和初始值比传统的非扰动法提供了更大的自由度。在 $q=0, q=1$ 时，有 $\phi(x;0) = u_0(x)$、$\phi(x;1) = u(x)$。因此，当嵌入参数 q 由 0 增长到 1，零阶形变方程的解

$\phi(x;q)$ 由初始解 $u_0(x)$ 变化到原始非线性微分方程 $N[u(x)]=0$ 的精确解 $u(x)$ 。这样的连续变化在拓扑中称为形变。

由于 $\phi(x;q)$ 也依赖于嵌入参数 $q\in[0,1]$ ，可将其扩展为关于 q 的 Maclaurin 级数，称为 Homotopy-Maclaurin 级数。

$$\phi(x;q)=u_0(x)+\sum_{n=1}^{+\infty}u_n(x)q^n \tag{2.29}$$

需要说明的是，选择辅助线性算子 L 和初始解 $u_0(x)$ 具有极大的自由度。如果选择恰当，则 Homotopy-Maclaurin 级数在 $q=1$ 处收敛，称为 Homotopy 级数解。

$$u(x)=u_0(x)+\sum_{n=1}^{+\infty}u_n(x) \tag{2.30}$$

该级数满足原始方程 $N[u(x)]=0$ 。这里的 $u_n(x)$ 受高阶形变方程控制，即

$$L[u_n(x)-\chi_n u_{n-1}(x)]=-\delta_{n-1}(x) \tag{2.31}$$

其中，

$$\chi_k=\begin{cases}1 & ,k\geqslant 2\\ 0 & ,k<2\end{cases}$$

$$\delta_k(x)=\left.\left\{\frac{1}{k!}\frac{\partial^k N[\phi(x',q)]}{\partial q^k}\right\}\right|_{q=0} \tag{2.32}$$

2.4.2 标准 HAM

廖世俊发现，早期的 HAM 并非总能确保一般的非线性方程组的逼近级数收敛[78,79]。为了克服这个缺陷，廖世俊引进一个非零辅助参数 c_0 来构造双参数方程簇。此时零阶形变方程为

$$(1-q)L[\phi(x',q)-u_0(x)]=c_0qN[\phi(x;q)],\ \ x\in\Omega,q\in[0,1] \quad (2.33)$$

对应的高阶形变方程为

$$L[u_n(x)-\chi_n u_{n-1}(x)]=c_0\delta_{n-1}(x) \quad (2.34)$$

这样，式(2.30)中的 Homotopy 级数解不仅依赖于物理变量 x，而且依赖于附加的辅助参数 c_0。从数学上来看，c_0 可以调节和控制收敛域[80]，尽管 c_0 没有任何物理意义。

虽然辅助参数 c_0 增加了“人为”的自由度，但却大大改善了早期 HAM，提供了一种方法以确保 Homotopy 级数解的收敛性。因此，c_0 被称为收敛控制参数。这是 HAM 的一个显著进步。

通过上面的介绍，可以发现，似乎更多的“人造”自由度可以大大提高使用同伦分析法获得较好的逼近解的概率。因此，廖世俊又提出了更一般的零阶形变方程[81]。

$$[1-\alpha(q)]L[\phi(x',q)-u_0(x)]=c_0\beta(q)N[\phi(x',q)],\ \ x\in\Omega,q\in[0,1] \quad (2.35)$$

其中，$\alpha(q),\beta(q)$ 称为形变函数，满足

$$\begin{cases}\alpha(0)=\beta(0)=0\\ \alpha(1)=\beta(1)=1\end{cases} \quad (2.36)$$

对应的泰勒级数为

$$\alpha(q)=\sum_{m=1}^{\infty}\alpha_m q^m\ ,\ \beta(q)=\sum_{m=1}^{\infty}\beta_m q^m \tag{2.37}$$

当$|q|\leqslant 1$，收敛。对应的高阶形变方程为

$$L\left[u_m(x)-\sum_{k=1}^{m-1}\alpha_k u_{m-k}(x)\right]=c_0\sum_{k=1}^{m}\beta_k\delta_{m-k}(x) \tag{2.38}$$

根据上面的介绍可见，由于 HAM 的优越性能，可以推广到求解偏微分方程。

2.4.3　优化 HAM

通过研究发现，存在一个有效区域R_c，对于任何$c_0\in R_c$都能获得一个收敛级数解。可以大致绘制出关于c_0的一条曲线。但这样的曲线并不能告诉哪个收敛控制参数c_0是最好的，即对应于最快的收敛级数。

Yabushita 等人[82]使用 HAM 来求解双耦合非线性常微分方程，并提出一种新的优化方法，通过极小化给定方程的残差平方找出两个最优化收敛控制参数。即令

$$E_m=\int_{\Omega}\left\{N\left[\sum_{n=1}^{m}u_n(x)\right]\right\}^2\mathrm{d}\Omega$$

为方程$N(u)=0$的第m阶逼近的残差平方，在整个Ω域进行积分。理论上，如果$E_m\to 0$，则$\sum_{n=1}^{\infty}u_n(x)$是$N(u)=0$的一个级数解。

因此，如果仅存在一个收敛控制参数c_0，则R_c定义为

$$R_c = \left\{ c_0 \mid \lim_{m \to +\infty} E_m(c_0) = 0 \right\} \tag{2.39}$$

另外，在一个给定阶次的逼近，E_m相当于优化的逼近。因此，E_m关于c_0的曲线不仅指明了R_c，而且指出了最优化的c_0，即对应于E_m的极小值。需要注意的是，对任何阶次的逼近，都能获得一个E_m，即使精确解未知。

后来，Akyildiz 等人[83]通过极小化E_m获得一个收敛控制参数，并发现 Homotopy 级数解的收敛非常快。Marinca 等人[84]在式(2.35)所示的零阶形变方程中组合了c_0和$\beta(q)$，作为一个函数$\breve{\beta}(q)$，有$\breve{\beta}(0)=0, \breve{\beta}(1)\neq 1$，并构成一个方程簇为

$$(1-q)L[\phi(x',q)-u_0(x)] = \breve{\beta}(q)N[\phi(x',q)], \quad x \in \Omega, q \in [0,1] \tag{2.40}$$

泰勒级数为

$$\breve{\beta}(q) = \sum_{n=1}^{\infty} c_n q^n$$

收敛在$q=1$。

式(2.40)是式(2.35)的一个特例，即令

$$\alpha(q) = q, \beta(q) = \frac{1}{c_0}\sum_{n=1}^{\infty} c_n q_n = \frac{\breve{\beta}(q)}{c_0}, c_0 = \sum_{n=1}^{\infty} c_n \neq 0 \tag{2.41}$$

称为优化同伦渐进分析法。

2.5 半定规划

我们通常所说的线性规划(Linear Programming，LP)是指极小化或极大化含 n 个变量的线性函数，要求服从线性的等式和不等式约束，可写成下述方程形式。

$$\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{Ax}=\boldsymbol{b}, \boldsymbol{x} \geqslant \boldsymbol{0}$$

其中，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 为含 n 个变量的列向量，$\boldsymbol{c}=(c_1,c_2,\cdots,c_n)$ 是目标函数向量，$\boldsymbol{b}=(b_1,b_2,\cdots,b_m)$ 是右手边。

$\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 是约束矩阵，$\boldsymbol{0}$ 是适当维的零向量。

2.5.1　线性规划的一种推广

为了给出半定规划的定义，首先将线性规划在以下 4 个方面进行推广。

(1) 将 $\boldsymbol{x}$ 的向量空间 $\mathbb{R}^n$ 替换为另一个向量空间 $\mathrm{SYM}_n=\{\boldsymbol{X}\in\mathbb{R}^{n\times n}:x_{ij}=x_{ji},1\leqslant i<j\leqslant n\}$，即 $\boldsymbol{X}$ 为 $n\times n$ 的对称矩阵的空间。

(2) 将约束矩阵 $\boldsymbol{A}$ 替换为一个线性映射 $\boldsymbol{A}:\mathrm{SYM}_n\to\mathbb{R}^m$。

(3) $\mathbb{R}^n$ 上的标准的标量积 $<\boldsymbol{x},\boldsymbol{y}>=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$ 替换为 SYM_n 上的标量积

$$\boldsymbol{XY}=\sum_{i=1}^{n}\sum_{j=1}^{n}x_{ij}y_{ij}$$

也可以记为 $\boldsymbol{XY}=\mathrm{Tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y})$。对于方阵 $\boldsymbol{M}$， $\mathrm{Tr}(\boldsymbol{M})$ 为 $\boldsymbol{M}$ 的对角元素之和。

(4) 将约束 $\boldsymbol{x}\geqslant\boldsymbol{0}$ 替换为 $\boldsymbol{X}\succeq 0$。 $\boldsymbol{X}\succeq 0$ 表示 $\boldsymbol{X}$ 是半正定的。

2.5.2 正定矩阵与半正定矩阵

定义 2.6(正定矩阵)

一个矩阵 $\boldsymbol{M}$，如果对于所有的 $\boldsymbol{x}\neq 0$，有 $\boldsymbol{x}^{\mathrm{T}}M\boldsymbol{x}>0$，则 $\boldsymbol{M}$ 称为正定矩阵。

定义 2.7(半正定矩阵)

一个半正定矩阵 $\boldsymbol{M}$ 是一个实矩阵，是对称的，且所有特征值为非负的。记为 $\boldsymbol{M}\succeq 0$。

一个对称实矩阵仅有实特征值，因此较易理解非负特征值的条件。

定理 2.3

若 $\boldsymbol{M}\in\mathrm{SYM}_n$，下列说法是等价的。

(1) $\boldsymbol{M}$ 是半正定，对于所有的 $\boldsymbol{x}\in\mathbb{R}^n$，有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{x}\geqslant 0$。

(2) $\boldsymbol{M}$ 的所有特征值非负。

(3) 存在一个矩阵 $\boldsymbol{U}\in\mathbb{R}^{n\times n}$，使得 $\boldsymbol{M}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$。

证明：

(1) 证明如果 $\boldsymbol{M}\succeq 0$，则 $\boldsymbol{M}$ 的所有特征值非负。

根据定义， $\boldsymbol{M}$ 的特征向量 $\boldsymbol{\omega}$ 对于某特征值 λ，要求满足

$$M\boldsymbol{\omega}=\lambda\boldsymbol{\omega} \tag{2.42}$$

当 $\boldsymbol{M}$ 是对称的，所有特征值是实值。

将式(2.42)两边乘以 $\boldsymbol{\omega}^{\mathrm{T}}$，得到

$$\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{\omega}=\lambda\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} \tag{2.43}$$

根据半正定矩阵的定义，式(2.43)的左端是非负的，而右端 $\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}\geqslant 0$，所以 $\lambda\geqslant 0$。

(2) 证明如果 $\boldsymbol{M}$ 的所有特征值非负，则 $\boldsymbol{M}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$，其中 $\boldsymbol{U}\in\mathbb{R}^{m\times n}$。

设 $\boldsymbol{\Lambda}$ 为包含 $\boldsymbol{M}$ 的所有特征值的对角矩阵，即

$$\boldsymbol{\Lambda}=\begin{bmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{bmatrix} \tag{2.44}$$

令 $\boldsymbol{W}$ 为相当于向量 $[\boldsymbol{\omega}_1 \quad \boldsymbol{\omega}_2 \quad \cdots \quad \boldsymbol{\omega}_n]$ 的矩阵，由于特征向量是互相正交的，即

$$\boldsymbol{\omega}_i^{\mathrm{T}}\boldsymbol{\omega}_j=\begin{cases}1 & ,i=j\\ 0 & ,i\neq j\end{cases} \tag{2.45}$$

则 $\boldsymbol{W}$ 是一个正交矩阵。

给定 $\boldsymbol{W}$ 和 $\boldsymbol{\Lambda}$，得到特征值的矩阵表示为 $\boldsymbol{MW}=\boldsymbol{W\Lambda}$。两边乘以 $\boldsymbol{W}^{\mathrm{T}}$，得到

$$\boldsymbol{MWW}^{\mathrm{T}}=\boldsymbol{W\Lambda W}^{\mathrm{T}} \tag{2.46}$$

由于 $\boldsymbol{W}$ 是一个正交矩阵，即 $\boldsymbol{WW}^{\mathrm{T}}=\boldsymbol{I}$，这样，式(2.46)的左端就等于 $\boldsymbol{M}$。又由于所有特征值是非负的，则将 $\boldsymbol{\Lambda}$ 分解为 $\boldsymbol{\Lambda}^{1/2}(\boldsymbol{\Lambda}^{1/2})^{\mathrm{T}}$，其中 $\boldsymbol{\Lambda}^{1/2}$ 定义如下。

$$\boldsymbol{\Lambda}^{1/2}=\begin{bmatrix}(\lambda_1)^{1/2} & & 0\\ & \ddots & \\ 0 & & (\lambda_n)^{1/2}\end{bmatrix} \tag{2.47}$$

这样，$\boldsymbol{W\Lambda W}^{\mathrm{T}}=\boldsymbol{W\Lambda}^{1/2}(\boldsymbol{\Lambda}^{1/2})^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}=\boldsymbol{W\Lambda}^{1/2}(\boldsymbol{W\Lambda}^{1/2})^{\mathrm{T}}$。令 $\boldsymbol{U}^{\mathrm{T}}=\boldsymbol{W\Lambda}^{1/2}$，则式(2.47)等价于 $\boldsymbol{M}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$。

(3) 证明如果 $\boldsymbol{M}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$，其中 $\boldsymbol{U}\in\mathbb{R}^{m\times n}$，则 $\boldsymbol{M}\succeq 0$。

根据 $\boldsymbol{M}$ 是半正定矩阵的定义，有

$$\boldsymbol{M}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$$

$$\Leftrightarrow \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Mx}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{Ux}\text{ , }\boldsymbol{x}\in\mathbb{R}^n$$

$$\Leftrightarrow \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Mx}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}\text{ , }\boldsymbol{y}\in\boldsymbol{Ux}$$

$$\Leftrightarrow \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Mx}\geqslant 0\text{ , }\because\ \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}\geqslant 0$$

□

根据定理 2.3 第(3)条，一个半正定矩阵可以分解为 $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$，令

$$\boldsymbol{U}=\begin{bmatrix}u_1 & u_2 & \cdots & u_n\end{bmatrix}\text{ , }u_i\in\mathbb{R}^m$$

则 A_{ij} 就是 C_i,C_j 的点积。因此，得到下面的推论。

定义 2.8(半正定矩阵集合)

PSD_n 是所有 $n\times n$ 半正定矩阵的集合，显然，有 $\mathrm{PSD}_n\subseteq\mathrm{SYM}_n$。

2.5.3　Cholesky 分解

对于一个半定正矩阵 $\boldsymbol{M}$，求解满足 $\boldsymbol{M}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$ 的矩阵 $\boldsymbol{U}$。这个过程称为 Cholesky 分解。如果 $\boldsymbol{M}=(\alpha)\in\mathbb{R}^{1\times1}$，则设 $\boldsymbol{U}=(\sqrt{\alpha})$，其中 $\alpha\geqslant 0$。若 $\boldsymbol{M}$ 对称的，则有

$$\boldsymbol{M}=\begin{pmatrix}\alpha & \boldsymbol{q}^{\mathrm{T}}\\ \boldsymbol{q} & \boldsymbol{N}\end{pmatrix} \tag{2.48}$$

则 $\alpha=\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{M}\boldsymbol{e}_1$，其中，$\boldsymbol{e}_i$ 为单位向量。有两种情况需要特殊考察。如果 $\alpha>0$，则有

$$M=\begin{pmatrix}\sqrt{\alpha} & \mathbf{0}^{\mathrm{T}}\\ \dfrac{1}{\sqrt{\alpha}}\boldsymbol{q} & I_{n-1}\end{pmatrix}\begin{pmatrix}1 & \mathbf{0}^{\mathrm{T}}\\ \mathbf{0} & N-\dfrac{1}{\alpha}\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}\sqrt{\alpha} & \dfrac{1}{\sqrt{\alpha}}\boldsymbol{q}^{\mathrm{T}}\\ \mathbf{0} & I_{n-1}\end{pmatrix} \tag{2.49}$$

矩阵 $\boldsymbol{N}-\dfrac{1}{\alpha}\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}}$ 也是半正定的，可以循环计算 Cholesky 分解。

$$\boldsymbol{N}-\frac{1}{\alpha}\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}}=\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V}$$

通过基本计算，得到

$$\boldsymbol{U}=\begin{pmatrix}\sqrt{\alpha} & \dfrac{1}{\sqrt{\alpha}}\boldsymbol{q}^{\mathrm{T}}\\ \mathbf{0} & V\end{pmatrix}$$

满足 $\boldsymbol{M}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$，可以找到 $\boldsymbol{M}$ 的一个 Cholesky 分解。

在其他的情况($\alpha=0$)，有 $\boldsymbol{q}=\mathbf{0}$。矩阵 $\boldsymbol{N}$ 是半正定的，可以计算

$\boldsymbol{N}=\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V}$。令

$$\boldsymbol{U}=\begin{pmatrix} 0 & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{0} & \boldsymbol{V} \end{pmatrix}$$

则给出 $\boldsymbol{M}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$。

2.5.4 半定规划

下面给出半定规划(Semi-Definite Programming，SDP)的形式化定义。

定义 2.9(半定规划)

一个半定规划描述了下面的最优化问题。

$$\max\sum_{i,j=1}^{n}c_{ij}x_{ij} \quad \text{s.t.} \quad \sum_{i,j=1}^{n}a_{ijk}x_{ij}=b_k\ ,\ k=1,\cdots,m \qquad (2.50)$$
$$X\succeq 0$$

其中，$x_{ij},1\leqslant i,j\leqslant n$ 是 n^2 个变量，满足对称性约束 $x_{ji}=x_{ij}$。而 c_{ij},a_{ijk},b_k 是实系数，且

$$\boldsymbol{X}=(x_{ij})_{i,j=1}^{n}\in \mathrm{SYM}_n$$

也可以写为下述更紧凑的形式。

$$\begin{aligned} &\max \boldsymbol{C}\boldsymbol{X} \\ &\text{s.t.}\ \ \boldsymbol{A}_i\boldsymbol{X}=b_i, i=1,2,\cdots,m \\ &\qquad \boldsymbol{X}\succeq 0 \end{aligned} \qquad (2.51)$$

其中，$\boldsymbol{C}=(c_{ij})_{i,j=1}^{n}$ 是表示目标函数的一个矩阵，且

$\boldsymbol{A}_k=(a_{ijk})_{i,j=1}^n$，$k=1,2,\cdots,m$，线性约束 $\boldsymbol{A}_i\boldsymbol{X}=b_i,i=1,2,\cdots,m$ 可以写成 $\boldsymbol{A}(\boldsymbol{X})=\boldsymbol{b}$，其中，$\boldsymbol{b}=(b_1,\cdots,b_m)$。$\boldsymbol{A}:\mathrm{SYM}_n\mapsto\mathbb{R}^m$ 是一个线性映射。如果 SDP 存在一个可行解 $\tilde{\boldsymbol{X}}\in\mathrm{SYM}_n$，满足 $\boldsymbol{A}(\tilde{\boldsymbol{X}})=\boldsymbol{b}$ 和 $\tilde{\boldsymbol{X}}\succeq 0$，则称 SDP 是可行的。一个可行的 SDP 的值为 $\sup\{\boldsymbol{CX}:\boldsymbol{A}(\boldsymbol{X})=\boldsymbol{b},\boldsymbol{X}\succeq 0\}$，当值为 ∞，SDP 规划称为无界的。否则，称为有界 SDP。

一个优化解是一个可行解 $\boldsymbol{X}^*$，对所有可行解 $\boldsymbol{X}$，满足 $\boldsymbol{CX}^*\geqslant\boldsymbol{CX}$。因此，如果有一个优化解，其对应的 SDP 值是有限的，且是可达的，则意味着上式的上确界是一个极大值。

2.5.5　非标准半定规划

SDP 并非总能写成式(2.51)的形式。除了式(2.51)中的线性等式约束外，还可能存在一些不等式约束，因此需要增加不属半正定矩阵 $\boldsymbol{X}$ 元素的实变量。

（1）首先，增加另外的未出现在 $\boldsymbol{X}$ 中的实变量 $x_1,x_2,\cdots,x_k$，将它们加入矩阵，形成 $\boldsymbol{X}'\in\mathrm{SYM}_{n+k}$。

$$\boldsymbol{X}'=\begin{pmatrix}\boldsymbol{X} & 0 & \cdots & 0\\ 0 & x_1 & 0 & 0\\ 0 & 0 & \ddots & 0\\ 0 & 0 & \cdots & x_k\end{pmatrix}\tag{2.52}$$

显然，当且仅当 $\boldsymbol{X}\succeq 0$ 和 $x_1,x_2,\cdots,x_k\geqslant 0$，有 $\boldsymbol{X}'\succeq 0$。

（2）为了消除不等式，可以增加非负的松弛变量，就像在线性规划一样。例如，一个不等式约束 $x_{23}+5x_{15}\leqslant 22$ 替换为一个等式约束 $x_{23}+5x_{15}+y=22$，其中的 y 是增加的非负实变量。

（3）无约束实变量 x_i 替换为一个差 $x_i^{'} - x_i^{''}$，其中 $x_i^{'}, x_i^{''}$ 为两个新的非负的实变量。

通过上述步骤，一个非标准式 SDP 可以转化为标准 SDP 形式。

2.5.6　半定规划可行区

类似于线性规划，一个 SDP 中的一个线性约束对应一个超平面，限制了 SDP 的可行区。而一个半定约束是非线性的，产生一个非纯平面。事实上，一个半定约束可表示为无限多的线性约束。下面来看图 2.19 所示的例子。

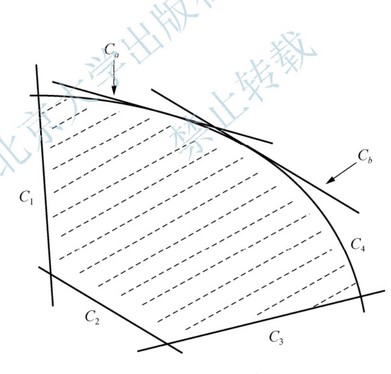


图 2.19　SDP 的可行区

其中，C_1,C_2,C_3 是三个线性约束，C_4 是非线性约束，相当于无穷多个线性约束。而 C_a,C_b 是这无穷多个线性约束中的两个实例。这些无穷多个非线性约束产生 C_4 的非纯平面。

一个 SDP 的优化解可能位于可行区的非纯平面上，这样，就是不合理的解。看下面的一个具体例子。基于下面的约束，极小化 x。

$$\begin{bmatrix} x & \sqrt{2} \\ \sqrt{2} & x \end{bmatrix} \succeq 0 \tag{2.53}$$

对于半正定矩阵，所有前主子式都是非负的。所谓前主子式，是指删除一个 $n \times n$ 矩阵中的最后 k 行和最后 k 列元素后而构成的子矩阵所对应的行列式。例如，式(2.53)的矩阵有两个前主子式，即

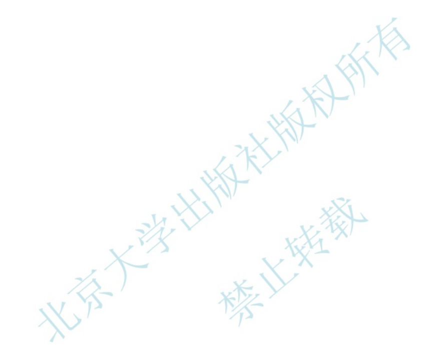
$$\begin{cases} m_1 = \left| [x] \right| \\ m_2 = \left| \begin{bmatrix} x & \sqrt{2} \\ \sqrt{2} & x \end{bmatrix} \right| = x^2 - 2 \end{cases} \tag{2.54}$$

这样，有 $x \geqslant 0$ 和 $x^2 - 2 \geqslant 0$。立即可以得到 x 的最小值为 $\sqrt{2}$。

2.6　本 章 小 结

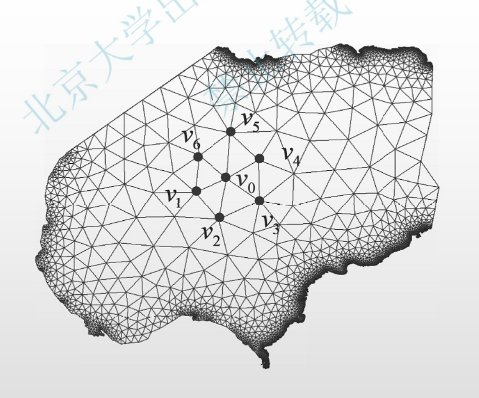
本章先扼要介绍了 3D 表面的形式化定义及两种基本类型，着重分析了平滑表面成为局部平面的必要条件。接着讨论了光照原理及针孔相机模型，重点探讨了 Lambertian 表面模型的基本条件，以及单光

源假设的合理性和必要性。在此基础之上进一步介绍了光学成像过程，为后面的三种 3D 重构提供知识支撑。最后介绍了同伦分析法和半定规划两种较流行的非线性优化方法，具体应用将在第 3、4、5 章详细讨论。



第 3 章

非刚体表面的 SFS 闭合解模式



非刚体表面在现实中普遍存在，如弯曲的柔性板材、压缩的球面、扭曲的脸部、动画片中的卡通对象等。基于单图像的非刚体表面重构严重约束不足，通用办法是引进形变模型，或者基于物理特性[85,86,87,88,89]，或者训练测试数据[36,90,91,92,93]，将表面形变表示为模型参数，再选择恰当的初始化模型，通过极小化一个基于图像的目标函数而获得解。由于这些目标函数可能具有多个局部极小值，因此严重依赖于初始解的选取。而 SFS 是迄今发展最完善的单图像表面重构方法，衍生出许多变异类型，广泛应用于计算机视觉的各个工程领域。因此，本章将 SFS 问题规划为多项式系统，应用多项式系统的相关特性，提出三种基于 SFS 重构 3D 表面的新方法，分别用于非刚体表面的 3D 重构。

3.1 引　言

由于投影成像过程丢失了深度信息，不同的形状可能产生同一个图像，因此，单幅图像的 3D 重构是一个非适定病态问题，通常需要附加额外的假定条件。

3.1.1 标准 SFS 模式

Horn 方法[38]假定表面遵守 Lambertian 反射特性(即漫反射)，如图 3.1 所示。

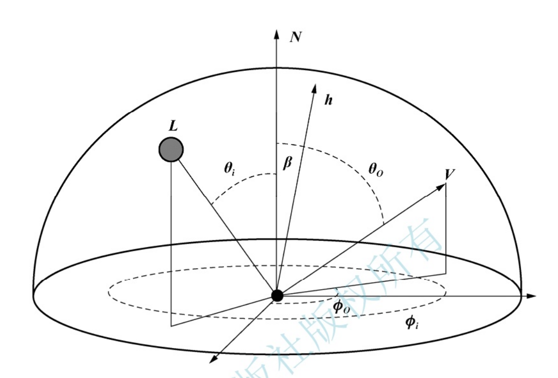


图 3.1　Lambertian 表面反射模型

其中，$\boldsymbol{L}=(a,b,c)$ 表示光源方向，$\boldsymbol{N}=(-p,-q,1)^{\mathrm{T}}$ 表示表面法向，$\boldsymbol{V}$ 表示视向。

Lambertian 反射遵守余弦定律。

$$I_{\mathrm{o}}=\rho \boldsymbol{N}\cdot \boldsymbol{L} I_{\mathrm{i}} \tag{3.1}$$

其中，I_{i}、I_{o} 分别表示输入光强(即光源强度)、输出光强(即观察到的图像亮度)；ρ 为反照率。

式(3.1)表明，输出光强仅与光源方向和表面法向的夹角余弦成比例，而不依赖于视向，即在所有方向观察表面感受到相等的图像亮度。

理论上讲，若已知 $\rho,\boldsymbol{L},I_{\mathrm{i}}$，可以根据 I_{o} 恢复表面法向 $\boldsymbol{N}$，这是表面重构的目标。这个求解过程类似于积分运算，所要求解的表面法向 $\boldsymbol{N}$ 相当于一个导数。此处困难的是，有一部分导数（$\boldsymbol{N}$）无法确定。因此，必须假设表面是连续的。

假设表面表示为$z(x,y)$，令

$$z_x = p = \frac{\partial z}{\partial x}\ ,\ z_y = q = \frac{\partial z}{\partial y} \tag{3.2}$$

表面上某点$(x,y,z(x,y))^{\mathrm{T}}$的切向分别为$\boldsymbol{t}_x = (1,0,z_x)^{\mathrm{T}}$、$\boldsymbol{t}_y = (1,0,z_y)^{\mathrm{T}}$，则该点的法向为

$$\boldsymbol{N} = \frac{\boldsymbol{t}_x \cdot \boldsymbol{t}_y}{\left\|\boldsymbol{t}_x \cdot \boldsymbol{t}_y\right\|} = \frac{1}{\sqrt{p^2 + q^2 + 1}}(-p,-q,1)^{\mathrm{T}} \tag{3.3}$$

因此，该点的图像强度为

$$I(x,y) = \rho \boldsymbol{N} \cdot \boldsymbol{L} I_{\mathrm{i}} = \rho \frac{-ap - bq + c}{\sqrt{1 + p^2 + q^2}} I_{\mathrm{i}} \tag{3.4}$$

对于$M \times N$的图像的每一个像素都有一个类似式(3.4)的方程，最终得到$M \times N$个方程。

一般地，图像中的最亮点具有指向光源方向的一个法向，因此，最亮点的图像强度为ρI_{i}。如果以ρI_{i}为比例因子，则式(3.4)简化为

$$I(x,y) \propto \boldsymbol{N} \cdot \boldsymbol{L} = \frac{1}{\sqrt{p^2 + q^2 + 1}} \tag{3.5}$$

式(3.5)表明，不管物体表面是否为 Lambertian 都可以得到上述函数的某个比例值。若存在几个理想光源，且在 Lambertian 表面的一个区域内都是可见的，则式(3.4)中的ρI_{i}就是这些光源的线性叠加。因此，很多文献中关于单光源的假设是有一定道理的。

若光源 $\boldsymbol{L}=(a,b,c)$ 已知，则式(3.4)仅含有两个未知量 p,q，而 p,q 来自于可积表面 $z(x,y)$，因此可利用表面的可积性约束。

$$\frac{\alpha^2 z}{\partial x\partial y}=\frac{\alpha^2 z}{\partial y\partial x} \Rightarrow \frac{\partial p}{\partial y}=\frac{\partial q}{\partial y} \tag{3.6}$$

这样，最后得到的系统含有 $2M\times N$ 个非线性微分方程。需要注意的是，如果含有噪声或深度不连续，则系统无解。

整理式(3.5)，可得

$$\sqrt{p^2+q^2}=\sqrt{\frac{1}{I^2}-1}=\|\nabla z\| \tag{3.7}$$

式(3.7)称为光程方程(也称短时距方程)，这种情况下，图像中的 attached 或 cast 阴影不可见。需要注意的是，$\boldsymbol{L}=(0,0,1)$ 时存在局部歧义性，因为对于任意解表面的每一个局部面片存在 in/out 倒置的歧义性。

Horn 基于 SFS 的方法虽然具有优美的数学形式，但其关于 Lambertian 反射率的假设对于多数实际对象并不适用，可能导致无解或多重歧义解。因此，有不少研究人员致力于研究 SFS 解的存在性和唯一性。Horn 等人[94]证明了两类情况 SFS 无解，一是光程方程仅含无边界解，即物理上不可能实现；二是光程方程根本无解。Oliensis[95]证明了 SFS 解的唯一性条件是存在边界约束。Saxberg[96]证明了在靠近泛奇异点处存在多达四个局部解。Okatani 和 Deguchi[97]指出，如果表面不存在二阶导数或不连续，则可能存在无穷多解。Deift 和 Sylvester[98]则进一步研究了 Lambertian 半球图像的歧义性，指出如果表面是类 C^2 型，则半球可克服 in/out 倒置歧义性。总的来说，尽管从

理论上讲 SFS 可能存在唯一解，或者说存在一定数量的解，但在实践中，SFS 问题仍然是一个严重非适定病态问题。

Zhang 等人[99]将 SFS 方法大致划分为四类：极小化法[100]、传播法、局部法以及线性解法。目前的大多数 SFS 方法或者基于极小化误差，或者致力于改进传播方法。不过，传播法严重依赖于边界和初始条件，或者从已知边界条件入手[101,102,103]，或者从奇异点入手[43,104,105]。而全局极小化方法通常从简单表面(如平面)开始[106]。

在很长一段时间内，对 SFS 的研究侧重于算法的普适性研究，近年来，通过假定某些约束条件来简化模型的研究重新得到了关注[107,108,109,110]。

3.1.2 非刚体 3D 表面重构

基于单图像的形变 3D 表面的重构问题由于其天然缺陷呈现严重约束不足，通用办法是引进形变模型，如超二次型[88]、三角形表面[91]、薄板样条[87]等，将表面形变表示为模型参数，首先初始化模型，然后极小化一个基于图像的目标函数。由于这些目标函数具有多个局部极小值，因此选取一个恰当的初始化至关重要，而这是相当困难的。近年来有些方法提出基于一个形状已知的参考图像寻找方法来逼近该形状[111,112]，而不需要形变的任何知识，但假设表面不可伸展，且邻近点之间的距离保持不变。这种假设对于诸如硬纸板等无弹性材质较合适，但对于弹性材料却不适用。模型分析虽然可以降低问题维数，但当用于检索形变对象时，不能捕捉复杂形变的物理现象，因为需要控制更复杂的非线性模型[89]。

传统的 SFS 和 SFT 通过强制表面平滑性，并假设表面是 Lambertian 类，或者统计出同质纹理分布来实现上述目标。这些方法的不足之处是，需要过多的假定条件，或者即使能够得到理想的逼近解，其迭代过程也需要较好的初始假设。而且这些方法不适用于弹性表面。有文献[113]提出组合使用 SFS 和 SFT，以克服解的歧义性，但是包含了多重迭代操作。

上述这些方法无论如何改进，总是需要模型初始化，然后极小化基于图像的目标函数，但这可能存在多个局部极小值。例如，在逐帧跟踪方案中，在某一帧中获得的形状参数虽然可以充当下一帧的初始值，但仍然需要指定第一帧的参数，而且如果跟踪失效，则无法完成重构。为了避免这种失败，必须寻求一种不依赖初始假定解即可恢复 3D 形状的新方案。有文献[93,111,112]讨论了这个问题，但是其都依赖于表面点间的距离保持恒定这样一个假定条件，这对于非弹性对象是合适的，而对于弹性对象则是无效的。本章所提出的方案对弹性或非弹性对象都适用，而且存在闭合式解。

3.1.3　基于多项式的 SFS 模式

我们知道，线搜索、同伦分析法、半定规划等非线性优化算法可用于多项式系统的求解[114]等。因此，一个很自然的想法是，如果能将 SFS 问题转化为一种多项式结构，则可使用这些成熟的算法来重构 3D 表面。

仍然以 Lambertian 模型为例，并令入射光强 $I_{\mathrm{i}}=1$，反照率 $\rho=1$，则根据式(3.4)有

$$I = \frac{-ap - bq + c}{\sqrt{1 + p^2 + q^2}} \in [0,1] \tag{3.8}$$

对式(3.8)两边取平方，经整理得到一个二次方程为

$$(1 + p^2 + q^2)I^2 - (-ap - bq + c)^2 = 0 \tag{3.9}$$

为避免 $I < 0$，增加一个约束条件为

$$-ap - bq + c \geqslant 0 \tag{3.10}$$

这样，SFS 问题归结为寻找一个表面 $z(x, y)$，其图像中每个像素的强度都满足式(3.9)和式(3.10)。

很明显，式(3.9)形式的多项式形式与式(3.8)的商形式相比，具有很多好处。首先，基于多面体的 SFS 多项式系统可用同伦分析法完全求解[115]；其次，若将 SDP 用于多项式系统可产生较理想的逼近解，而不依赖于初始值；第三，精确线搜索非常适用于多项式系统，使得 SFS 迭代算法具有较高的执行效率[116]。

3.2 构建同伦分析框架的 SFS 模式

针对 Horn 方法的计算复杂性，Lee 等人[117]提出了高效的迭代算法，Penna[107]则将其拓展到透视投影情况。Yang[118]、Shimodaira[119]等人则研究了存在多重解的可能性，将空间划分技术用于 SFS 求解，但当空间维数增加时，收敛速度将变得相当慢[120]。由于 SFS 可以转

化为多项式系统，可以应用同伦分析法，即从一个结构上相似且所有解已知的多项式系统开始，在从这个相似的多项式系统连续形变到所期望的多项式系统的过程中，跟踪解路径。同伦分析法可确保寻找到所有解，而不需要任何初始值。

为不失一般性，假设表面由三角形网构成，考察某顶点 v_0 及其 k 个按逆时针方向排列的相邻顶点 $v_1,\cdots,v_k$，如图 3.2 所示。



图 3.2　围绕一个顶点的一种三角剖分

为了简单化，将坐标系平移，使得 $v_0=(0,0,0)$。则一个三角形 (v_0,v_i,v_{i+1}) 的法向与 $v_i\cdot v_{i+1}=(y_i z_{i+1}-y_{i+1}z_i, x_{i+1}z_i-x_i z_{i+1}, x_i y_{i+1}-x_{i+1}y_i)$ 成比例。将前两个坐标值除以最后的那个坐标 $x_i y_{i+1}-x_{i+1}y_i$，得到 $v_i\cdot v_{i+1}\propto(-p_i,-q_i,1)$，其中 $-p_i,-q_i$ 分别为

$$-p_i=\frac{y_i z_{i+1}-y_{i+1}z_i}{x_i y_{i+1}-x_{i+1}y_i}\ ,\ -q_i=\frac{x_{i+1}z_i-x_i z_{i+1}}{x_i y_{i+1}-x_{i+1}y_i} \tag{3.11}$$

对于邻域范围内的每个三角形，将式(3.11)代入式(3.9)中，得到 k 个关于 $\boldsymbol{z}=(z_1,\cdots,z_k)$ 的二次方程，简记为 $f(\boldsymbol{z})=0$，至多存在 2^k 个解。

若已知 $f(z)=0$ 的解的大概情况(例如，已知位于某个已知点附近)，则可使用类似牛顿法来求解。若对目标解一无所知，如果随机地选一个初始点，则可能导致系统发散而无法产生目标解。这时候就可使用同伦分析法来求解。

所谓同伦分析法，就是从一个已知解的结构类似的简单函数 $g(z)$ 逐渐形变到目标函数 $f(z)$ 的过程。这时候需要构造一个新函数。

$$h'(z,t)=t\cdot f(z)+(1-t)\cdot g(z) \tag{3.12}$$

其中，$t\in[0,1]$ 称为同伦参数。$g(z)=0$ 为解已知的结构类似的方程。显然，当 $t=0$ 时，$h'(z,0)=g(z)$。同伦分析过程就是使 $t=0\to 1$ 逐步变化，直到 $t=1$ 时 $h'(z,1)=f(z)$，即表示找到了 $f(z)=0$ 的解。为不失一般性，这里考察复数域 $\mathbb{C}$。当用多面体同伦方法解一个多项式系统 $f(z)=0$，其中 $f(z)=(f_1(z),\cdots,f_n(z))\in\mathbb{C}^n$，$f_j(z)\in\mathbb{C}$，向量 $z=(z_1,z_2,\cdots,z_n)\in\mathbb{C}^n$。较多的文献资料已经充分证明，对于大多数多项式系统，多面体同伦方法足可以找到系统所有的孤立解。

为提高数值稳定性，令 $s\in(-\infty,0]\to t=\exp(s)\in[0,1]$，再令 $h(z,s)=h'(z,\exp(s))$，则有

$$h(z,-\infty)=\lim_{s\to-\infty}h(z,s)=h'(z,0) \tag{3.13}$$

如果选择一个足够大的正数 s^0，则 $h'(z,0)=0$ 的每个解 z^0 充当 $h(z,-s^0)=0$ 的一个精确解，因此，可以从 $(z^0,-s^0)$ 开始，跟踪同伦系统 $h(z,s)$ 的一条解曲线，如图 3.3 所示。

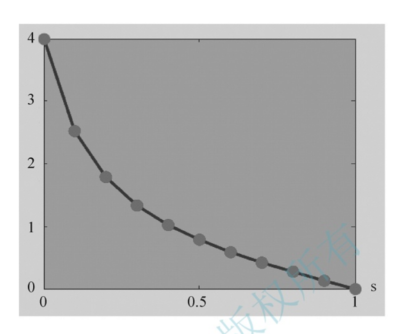


图 3.3　同伦分析法的解曲线

上述步骤可总结为表 3-1 所示的算法流程。

表 3-1　同伦方法求解表面的算法流程

步骤	内容
1	构建 $f(z)=0$ 方程，进而构造类似方程 $g(z)=0$，使得 $g(z)=0$ 易于求解
2	构造同伦函数 $h(z,s)=(h_1(z,s),h_2(z,s),\cdots,h_n(z,s))\in\mathbb{C}^n$，$(z,s)\in\mathbb{C}^n\times(-\infty,0]$，要求满足 $h(z,s)=\exp(s)f(z)+(1-\exp(s))g(z)$
3	令 $s=-s^0\to 0$，s^0 为一个足够大的正数，此时 $h(z,s^0)=g(z)$，解为 z^*
4	迭代下述过程，直到 $s=0$ (1) 令 $s'=s+\alpha$，$\alpha>0$ 为较小的迭代步长。 (2) 得到 $h(z,s')=\exp(s')f(z)+(1-\exp(s'))g(z)$。 (3) 应用牛顿法等求 $h(z,s')=0$ 的解 z'
5	验证 $s'=0$ 处 $h(z,s')=0$ 的解
6	返回 z'

在对围绕每一个顶点的所有可能构型都求解之后，需要进一步修剪解，即舍弃那些违反式(3.10)约束条件的构型，或邻点不存在兼容构

型的构型。其实质是解决一个约束满足问题，以识别所有的一致性全局解。由于某个顶点的一种构型及其邻点唯一确定了该顶点邻点的构型，因此，利用一个简单的指派传播和回溯算法即可找到所有全局解。

图 3.4 所示是基于同伦分析法的多面体的 SFS 效果。图 3.4(a)的表面含有 280 个大小各异的五边形局部面片；经预处理转化为含有 340 个三角形面片，如图 3.4(b)所示；再使用本节所述的同伦分析方法，重构出 3D 表面形状，如图 3.4(c)所示，该图已经过表面平滑化处理，显示出一种釉质感。



(a) 多面体

(b) 三角网格

(c) 重构出的 3D 表面

图 3.4　基于同伦分析法的一个多面体的 SFS 效果

3.3　使用线搜索优化迭代过程

线搜索法也称一维搜索法，是一种用于处理未知变量函数的优化迭代算法，是多变量优化的重要基础[121,122]。下面介绍如何使用线搜索方法处理 SFS 问题，以及如何优化 SFS 迭代过程。

3.3.1　建立线搜索基本框架

大多数 SFS 方法，如 Horn、Zhang[99]、Szeliski [123]等人着重于在一个栅格上极小化式(3.9)表示的平方和。设 $\boldsymbol{z}=(z_1,z_2,\cdots,z_n)^{\mathrm{T}}$ 为栅格上所有点的高度信息，令 $p_{ij}=z_{i+1,j}-z_{i,j}$， $q_{ij}=z_{i,j+1}-z_{i,j}$ 为表面某点 $v(i,j)$ 处在一个正交方向上的离散偏导数。

为避免指定边界条件，对于一幅 $M\times N$ 的图像，构造一个扩展栅格 $(M+1)\times(N+1)$，然后在这个栅格上求解 $\boldsymbol{z}$。这样，基于这个扩展的栅格，将式(3.9)表示的点 $v(i,j)$ 的多项式解模式 $(1+p^2+q^2)I^2-(-ap-bq+c)^2=0$ 转化为下述的该点邻域的解模式。

$$[z_{i,j}^2,z_{i+1,j}^2,z_{i,j+1}^2,z_{i,j}z_{i+1,j},z_{i,j}z_{i,j+1},z_{i+1,j}z_{i,j+1},z_{i,j},z_{i+1,j},z_{i,j+1},1]\cdot\boldsymbol{u}_{ij}=0 \tag{3.14}$$

其中，

$$\boldsymbol{u}_{ij}=[2I_{ij}^2,I_{ij}^2-a^2,I_{ij}^2-b^2,2a(a+b)-2I_{ij}^2,2b(a+b)-2I_{ij}^2,\\-2ab,-2c(a+b),2ac,2bc,I_{ij}^2-c^2]^{\mathrm{T}} \tag{3.15}$$

将式(3.14)和式(3.15)合并写成以下形式。

$$r_{ij}=\boldsymbol{z}^{\mathrm{T}}A_{ij}\boldsymbol{z}+e_{ij}^{\mathrm{T}}\boldsymbol{z}+h_{ij}=0 \tag{3.16}$$

其中，A_{ij} 为对称矩阵。

令 $r(\boldsymbol{z})=[r_{11},\cdots,r_{MN}]^{\mathrm{T}}$，则令

$$F(\boldsymbol{z})=\|\boldsymbol{r}\|^2 \tag{3.17}$$

其中，$\boldsymbol{r}$ 的雅可比矩阵的每一行表示为 $J_{ij}=2\boldsymbol{z}^{\mathrm{T}}A_{ij}+e_{ij}^{\mathrm{T}}$，对应的梯度表示为 $\nabla F=2\boldsymbol{J}^{\mathrm{T}}\boldsymbol{r}$。由于函数 $F(\boldsymbol{z})$ 是可微的，式(3.17)的优化问题可表示为

$$(P):\quad \min F(\boldsymbol{z}) \quad , \quad \text{s.t.} \quad \boldsymbol{z}\in\mathbb{R}^n \tag{3.18}$$

设一个起点 z_0，再通过迭代过程生成一个点的序列 $\{z_k\}_{k=0}^{\infty}$。当预设的某条件满足时，迭代过程终止。表 3-2 总结了上述算法的基本流程，输入：初始解 z_0。

表 3-2　线搜索优化算法

步骤	内容
1	For　$k=0,1,\cdots$ If　z_k 是优化的　Then 终止迭代。 Else (1) 确定一个搜索方向 d_k； (2) 确定一个步长 $\lambda_k>0$； (3) 计算新的解：$z_{k+1}=z_k+\lambda_k d_k$
2	返回结果

1. 确定搜索方向

可以简单地将 d_k 设为 $F(z)$ 在 z_k 处的下降方向，即

$$F(z_k+\lambda d_k)<F(z_k)\ ,\ \forall\lambda\in(0,\varepsilon]\ ,\ \varepsilon>0 \tag{3.19}$$

由于 $F(z)$ 是可微的，每当 $\nabla F(z_k)\neq 0$，满足 $\nabla F(z_k)^{\mathrm{T}}d_k<0$ 的任何 d_k 为一个下降方向。因此，可令 $d_k=-\boldsymbol{D}_k\nabla F(z_k)$，其中 $\boldsymbol{D}_k=H(z_k)^{-1}$ 为一个正定对称矩阵，其中 $H(z_k)=\nabla^2 F(z_k)$ 也为正定。

2. 选择迭代步长

一旦确定了 d_k，适当的迭代步长 λ_k 可以解决下面的一维优化问题。

$$\min_{\lambda\geqslant 0}F(z_k+\lambda d_k)$$

通过求解下面的一维优化问题，选定一个步长。

$$\lambda_k = \arg\min_{\lambda} \theta(\lambda) \triangleq \arg\min_{\lambda} F(\bar{z} + \lambda d) \tag{3.20}$$

因此假定 $F(z)$ 是可微凸函数，可尝试求解下式。

$$\bar{\lambda} = \arg\min_{\lambda>0} F(\bar{z} + \lambda\bar{d}) \tag{3.21}$$

此时的 $F(\bar{z} + \lambda\bar{d}) = \arg\min_{\lambda>0} F(\bar{z} + \lambda\bar{d})$，其中 $\bar{z}$ 为当前迭代，$\bar{d}$ 为当前搜索方向。

令 $\theta(\lambda) = F(\bar{z} + \lambda\bar{d})$，由于 $\theta(\lambda)$ 是凸函数，因此上述问题变为：

$$\bar{\lambda} = \arg\min_{\lambda>0} \theta(\lambda) \tag{3.22}$$

由于 $\theta(\lambda)$ 是凸函数，目标是找到一个 λ 使得 $\theta'(\bar{\lambda}) = 0$。由于 $\theta'(\lambda) = \nabla F(\bar{z} + \lambda\bar{d})$，$d$ 是下降方向，所以 $\theta'(0)<0$。假设已知一个满足 $\theta'(\hat{\lambda}) > 0$ 的 $\hat{\lambda} > 0$，使用表 3-3 所示的对分查找算法求解方程 $\theta'(\lambda) \approx 0$。

表 3-3　求解 $\theta'(\lambda) \approx 0$ 方程

步骤	内容
1	令 $k = 0, \lambda_l = 0, \lambda_u = \hat{\lambda}$
2	$\tilde{\lambda} = \dfrac{\lambda_u + \lambda_l}{2}$，并计算 $\theta'(\tilde{\lambda})$ (1) If $\theta'(\tilde{\lambda}) > 0$，令 $\lambda_u = \tilde{\lambda}, k = k+1$； (2) If $\theta'(\tilde{\lambda}) < 0$，令 $\lambda_l = \tilde{\lambda}, k = k+1$； (3) If $\theta'(\tilde{\lambda}) = 0$，终止
3	返回结果

至此得到一个关于 λ 的单变量四次函数，其全局极小化涉及解一个关于 λ 的三次方程。由于精确线搜索通常仅需要少量的函数求解，迭代计算量要比非精确线搜索少得多，而且会避免陷入一个低劣的局部极小化过程。为了更好地发挥线搜索的优势，下面介绍一种新的平滑项，以用于加宽收敛域、抑制波动。

3.3.2　正则化处理

根据式(3.6)，最简单也是最常用的平滑项(正则化)是对表面二阶导数的平方进行积分，目标是将表面不平坦的边缘处或光度显著变化的区域变得平坦化。有文献[124,125]建议通过亮度变化的一种度量，以降低平滑项的权重，增大平滑项可将亮度变化的表面区域变得弯曲，但不能呈现明显的折叠状；减小平滑项又可允许表面存在轻微的振荡，以增加表面的真实感。

首先考察两个法向分别为 $(-p_1,-q_1,1)$, $(-p_2,-q_2,1)$ ，亮度分别为 I_1,I_2 的相邻像素。两个法向之间的夹角满足式(3.23)。

$$\cos(\theta)=\cos(\arccos(I_1)-\arccos(I_2))=I_1\cdot I_2+\sqrt{1-I_1^2}\cdot\sqrt{1-I_2^2} \tag{3.23}$$

对于平滑表面上的比较邻近的像素，或者对于具有较大光亮度差的像素，要求法向之间的夹角接近于最小角 $\theta_{\min}$ ，即

$$\cos(\theta)\approx\frac{p_1p_2+q_1q_2+1}{\sqrt{1+p_1^2+q_1^2}\sqrt{1+p_2^2+q_2^2}}=\cos(\theta_{\min}) \tag{3.24}$$

代入式(3.8)，得到一个关于 z 的四次方程。

$$S(z)=\sum[(p_1p_2+q_1q_2+1)I_1I_2-\cos(\theta)(-ap_1-bq_1+c)(-ap_2-bq_2+c)]^2 \tag{3.25}$$

这样，上述问题归结为下面的最优化模式。

$$\arg\min_{z\in\mathbb{R}^n} F(z)+\lambda S(z) \tag{3.26}$$

图 3.5 分别通过人像与花瓶的重构效果验证了上述算法的有效性。

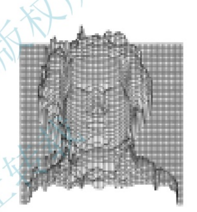
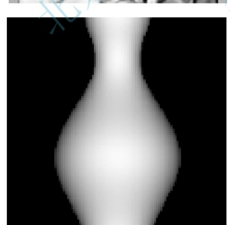
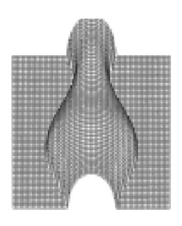





图 3.5　基于线搜索法的 SFS 效果图

3.4　将 SDP 推广到 SFS 问题

上述线搜索 SFS 需要预设一个初始值 z_0，而这个初始值若设置不当将会导致不理想的局部极小值，这显然不是我们所期望的。因此，需要寻找一种完全不同的方法以确保收敛于一个全局极小值。近年来，SDP 凸松弛方法被广泛用于解决基于距离和角度测量的几何问题[126,127]。一个很自然的想法是，能否将 SFS 问题设计成为类似的几何问题？答案是肯定的。基本思想是，引进一个新变量来表示变量之积，使得非凸项转化为线性，再增加一个半正定约束以描述所有变量之间的关系，这些半正定约束定义在单位三角形上。

对于表面上的每个点 (i,j)，使用变量的一个扩展集 x_{ij}^{rst}，理想情况下，x_{ij}^{rst} 相当于一个单项式 $(z_{ij})^r\cdot(z_{i+1,j})^s\cdot(z_{i,j+1})^t$，其中 r,s,t 为非负整数，满足 $r+s+t\leqslant 2d$，d 称为松弛阶次。这样，$x_{ij}^{k00},x_{i-1,j}^{0k0},x_{i,j-1}^{00k}$ 指向同一个相当于 $(z_{i,j})^k$ 的变量。当 $k=1$ 时，就表示点 (i,j) 处的表面高度，x_{ij}^{000} 代表常量 1。

下面将这些变量排列进动差矩阵。令 $(r_1,s_1,t_1)=(0,0,0)$，$(r_2,s_2,t_2),\cdots,(r_D,s_D,t_D)$ 是所有 r,s,t 的一个枚举，且满足式(3.27)。

$$\begin{cases} r+s+t \leqslant d \\ D = \begin{pmatrix} d+3 \\ 3 \end{pmatrix} = O(d^3) \end{cases} \tag{3.27}$$

对于每一个图像像素，建立一种新的约束条件，即式(3.28)表示的矩阵为一个 $D \times D$ 的、秩为 1 的半正定矩阵。

$$\left[1, x_{ij}^{r_2 s_2 t_2}, \cdots, x_{ij}^{r_D s_D t_D}\right]^{\mathrm{T}} \cdot \left[1, x_{ij}^{r_2 s_2 t_2}, \cdots, x_{ij}^{r_D s_D t_D}\right] \tag{3.28}$$

由于这个条件是非凸的，因此需要使用一个凸松弛。对于每个像素，一个半正定约束定义在一个 $D \times D$ 的动差矩阵上，该矩阵的元素是自由变量。

$$\left[\boldsymbol{M}_{ij}\right]_{m,n} = x_{ij}^{r_m + r_n, s_m + s_n, t_m + t_n}, \quad \left[\boldsymbol{M}_{ij}\right]_{1,1} = 1 \tag{3.29}$$

理想情况下，式(3.29)中的 $\boldsymbol{M}_{ij}$ 即为式(3.28)。

对于一个未被阴影笼罩的像素，上述 SDP 松弛可表示为式(3.30)。

$$\min\left(\sum_{ij} \operatorname{trace}(\boldsymbol{M}_{ij}) + G \cdot \varepsilon\right) \tag{3.30}$$

式(3.30)服从下面的四个条件。

$$(a+b)x_{ij}^{100} - a x_{ij}^{010} - b x_{ij}^{001} + c \geqslant 0 \tag{3.31}$$

$$\begin{aligned} -\varepsilon \leqslant [&x_{ij}^{r+2,s,t}, x_{ij}^{r,s+2,t}, x_{ij}^{r,s,t+2}, x_{ij}^{r+1,s+1,t}, x_{ij}^{r+1,s,t+1}, x_{ij}^{r,s+1,t+1}, x_{ij}^{r+1,s,t}, \\ &x_{ij}^{r,s+1,t}, x_{ij}^{r,s,t+1}, x_{ij}^{r,s,t}] \cdot u_{ij} \leqslant \varepsilon \end{aligned} \tag{3.32}$$

$$\sum_{ij} x_{ij}^{100} = 0 \tag{3.33}$$

$$\boldsymbol{M}_{ij} \succeq 0, \quad \left[\boldsymbol{M}_{ij}\right]_{1,1} = 1 \tag{3.34}$$

在上面的系统中，式(3.31)是将式(3.10)的约束条件用 p,q 离散差形式重新书写。而式(3.32)是将式(3.14)乘以单项式 $(z_{ij})^r \cdot (z_{i+1,j})^s \cdot (z_{i,j+1})^t$ 得到的，其中所有非负整数 r,s,t，满足 $r+s+t \leqslant 2d-2$；u_{ij} 是类似式(3.15)中的常向量；ε 是一个松弛变量。条件式(3.33)用于校准正投影中的全局深度歧义性。式(3.34)是来自于式(3.28)的半正定约束。式(3.30)的目标是极小化轨迹与 G 倍松弛变量的和，其中的 G 是所有动差矩阵中对角元素的总数。对于某个具体问题来说，G 是一个常量。理想地，矩阵 $\boldsymbol{M}_{ij}$ 的秩为 1。轨迹极小化通常用于逼近矩阵秩的极小值[128]。

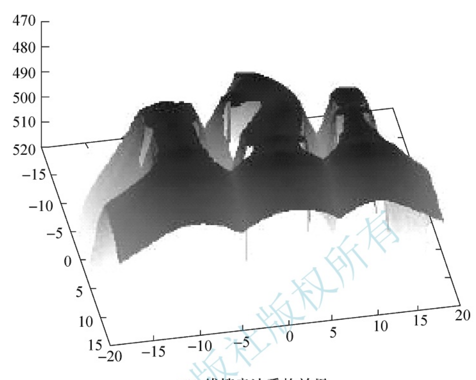
可利用上一章所述的 SDP 解法在多项式时间内获得结果 $\boldsymbol{M}_{ij}$，即式(3.28)，而不需要任何初始假定。

下面再从式(3.29)中的 $\boldsymbol{M}_{ij}$ 提取最终解。由于问题是凸性的，矩阵 $\boldsymbol{M}_{ij}$ 可能是解的凸组合。例如，当 $\boldsymbol{L}=(0,0,1)$，解 $z(x,y)$ 和 $-z(x,y)$ 这一个组合使得 $x_{ij}^{100}=0$。Waki 等人[137]提出一种伴随随机扰动的解提取过程，将边界条件加于偏微分方程上[130]。Henrion 等人[131]探讨了基于稠密松弛的解提取方法。图 3.6 给出了线搜索法和 SDP 法的表面重构效果对比。

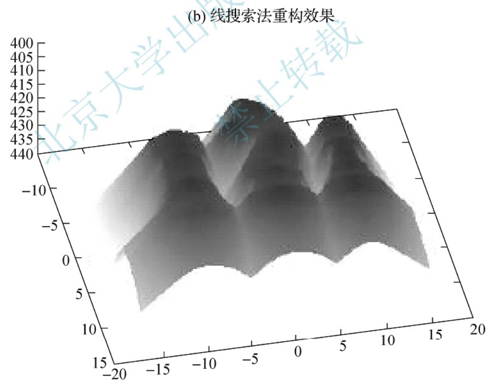


(a) 原图

图 3.6　线搜索法和 SDP 法的重构效果



(b) 线搜索法重构效果



(c) SDP法重构效果

图 3.6　线搜索法和 SDP 法的重构效果(续)

图 3.6 中的线搜索法的初始状态是 $z = 0$。使用同一个光源，松弛阶次 $d = 2$，从低分辨率图像得到 SDP 结果。图 3.7 是本章所述三种方法的重构效果的简单比较。

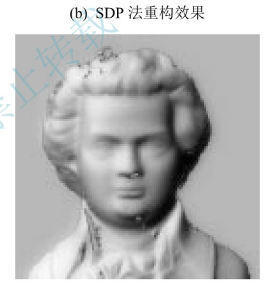


(a) 原图

(b) SDP 法重构效果



(c) 线搜索法重构效果



(d) 同伦分析法重构效果

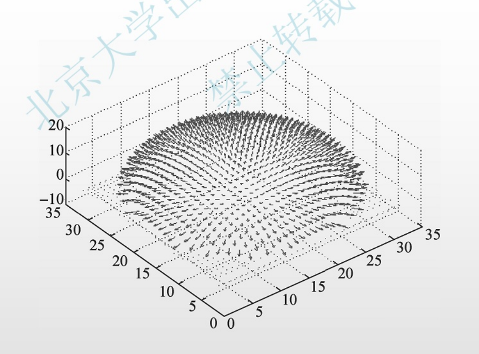
图 3.7　三种方法的 SFS 重构效果比较

3.5　本 章 小 结

经典 SFS 方法的理论依据是 $I_o = \rho \boldsymbol{N} \cdot \boldsymbol{L} I_i$，若已知光源方向 $\boldsymbol{L}$、入射光强 I_i 以及表面反照率 ρ，即可根据观察到的图像强度 I_o 恢复出表面法向 $\boldsymbol{N}$，再对 $\boldsymbol{N}$ 进行积分运算重构出 3D 表面形状。但如果表面不连续即存在 $\boldsymbol{N}$ 无法确定的表面区域，将导致 SFS 无解或无穷多解。另外，现有的 SFS 方法多基于迭代运算，需要假定一个恰当的初始解，若设置不当将导致最终解收敛于局部较小值或产生多重歧义解。且迭代过程复杂，计算量大，对于某些非凸项来说极可能导致无解或无穷多解。针对传统 SFS 的这种局限性，提出将 SFS 问题规划为多项式结构，再将成熟的同伦分析法、优化线搜索法和半定规划法等用于求解该多项式，设计了新的平滑项，使 SFS 摆脱了对初始假定的依赖性，优化了迭代过程，对非凸项实现了统一化的 SDP 凸松弛处理。这种将单图像的 3D 表面重构问题归入到一个成熟的理论框架体系的一种尝试，对 SFT、SFM 等重构方法的研究都具有重要的参考价值。实验结果表明，这种处理方案大大改善了 SFS 方法的高效性和普适性，对弹性表面和非弹性表面都有较好的重构效果。

第 4 章

离散歧义性的 SFT 解决方案



基于纹理的表面形状重构(SFT)方法主要根据表面上重复出现的单个纹理图案的畸变现象，或整个表面纹理的统计分布特性，来恢复表面法向信息。SFT 总体上类似于 SFS，但包含了更多的有效信息，如投影缩减现象等，但仍然需要一些假定约束条件。由于 SFT 是严重约束不足问题，常见的假定条件只能用于解决特定对象或特定类型的纹理表面。为此，本章主要讨论具有折叠歧义性的纹理图案，并将 SDP 法用于解决这种歧义性 SFT 问题。

4.1 引　言

纹理可分为确定性纹理和统计性纹理。确定性纹理是指重复显示的几何形状，如圆形、矩形及其他修饰图案。统计性纹理是指按固定概率变化的图案。由于 SFT 是约束不足问题，一般需要假定一些约束条件。例如，对于确定性纹理，一般假设相机已标定正投影环境，或假设纹理是同质的、各项同性的。如果观察到的一个表面片是平坦的，根据该面片与前视图之间的变换，可以计算出表面片的法向。而对于统计性纹理，一般假设以下一些前提条件。

(1) 正投影形成的纹理，即前视图。

(2) 整体表面或局部表面片是近似平坦的。

(3) 纹理元素是较小的线段，一般称为探针。

(4) 探针在整个表面和所有方向均匀分布。

使用局部法估计表面法向存在两个局限性：①必须知道表面片的

前视外观；②每个估计出的法向实际上是一个有歧义法向对，一是真实法向，二是穿越视向的镜像。Forsyth[49,132]和 Lobay[133]侧重估计呈现重复图案的前视纹理的法向，使用表面平滑来消除法向歧义性。而 Loh 和 Hartley[134]则通过透视效应来克服歧义性。基于单幅图像的 SFT 存在各种歧义性，其中较严重的是对称歧义性，如 in/out Necker 倒置歧义性(参见图 1.17)和二进制符号歧义性(参见图 1.16)。在这些约束不足情形下，根据不同的约束条件，表面可能有多种解释。SFT 方法需要从具有纹理的表面感知健壮的 3D 效果，如 4.1 图所示的是一些动物图案的纹理表面。



图 4.1　合成纹理表面

二进制歧义性较为简单，如表面图像的折叠歧义性，表面法向可从这种双折叠歧义性的纹理表面中估计出来。图 4.2 是一个特殊的二进制歧义性例子，重点考察了纹理元素之间的表面平滑性。

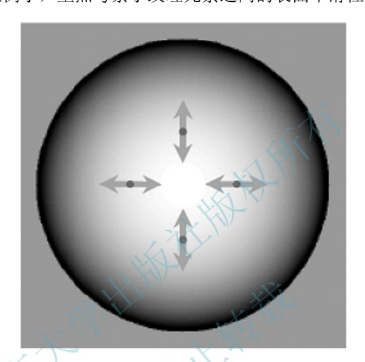
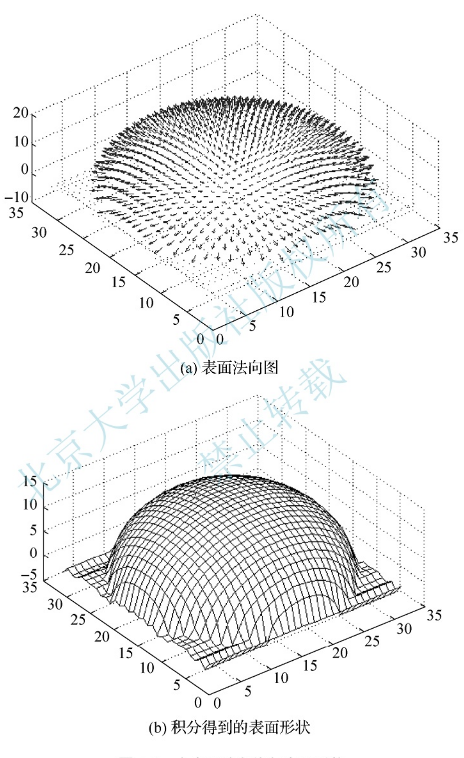


图 4.2　考察表面平滑性

图 4.2 考察了一个凹/凸球面。在四个位置上从纹理推断出形如 $(\pm a,0,1)^t$ 、 $(0,\pm b,1)^t$ 的有歧义性法向。每个法向与其最近邻法向之差的范数为 $\left\|(\pm a,0,1)^t-(0,\pm b,1)^t\right\|=\sqrt{a^2+b^2}$ 。若目标函数仅比较每条法向与其两条最近邻法向，则任意的交换分配都具有相同的价值。但是，这种依赖最近邻法向的特殊构型是无益的，因此须重视纹理元素间的表面平滑性。

在这种情况下描述图像表面带来两个挑战，一是必须在所有可能的法向中确定哪一个法向是正确的；二是必须将所选择的法向积分到一个候选表面上，如图 4.3 所示。



(a) 表面法向图

(b) 积分得到的表面形状

图 4.3　由表面法向恢复表面形状

在存在歧义性的情况下，这是一个难题，其不满足常用的一些诸如图割的优化方法的条件[135]。不过可以用半定规划法来实现逼近求解。SFT 问题的基本求解步骤如表 4-1 所示。

表 4-1　SFT 问题的求解步骤

步骤	内容
1	将表面表示为一个样条函数，即用一组连续变量控制形状，再用离散变量构成歧义性约束
2	定义一个二次目标函数来度量表面平滑度和约束满足度。消除连续变量，产生一个仅含离散变量的二次目标函数
3	使用 SDP 松弛法，将离散变量嵌入一个连续高维空间
4	采用圆整法整合离散变量，拟合一个 3D 形状

4.2　歧义性约束模式

从几何本质上看，SFT 非常类似于立体视觉和 SFM。为了将 SFT 归入到此类的多视图框架中，考察表面上具有同样纹理的邻近两个面片。这两个面片在单目视图下差异细微，得到单图像下的多视图结构。这样，很自然地存在一种两步骤求解模式：(1)从单图像计算纹理畸变；(2)解释这种纹理畸变以指示表面法向和 3D 形状。

因此，根据立体视觉的求解模式，实际应用中的表面重构问题可以写成以下的极小化形式[132]。

$$\underset{\boldsymbol{v},\boldsymbol{d}}{\arg\min}\left\|\boldsymbol{A}\boldsymbol{v}-\boldsymbol{B}\boldsymbol{d}\right\|^2 \tag{4.1}$$

其中，$\boldsymbol{A},\boldsymbol{B}$ 是矩阵；$\boldsymbol{d}\in\{-1,1\}^n$ 是离散判定变量的一个向量；$\boldsymbol{v}\in\mathbb{R}^m$ 是决定表面的连续变量的一个向量。将表面表示为一个样条函数，即为某基函数的一个线性组合。

$$z(x,y)=\sum_{i=1}^{m}\boldsymbol{b}_i(x,y)\boldsymbol{v}_i \tag{4.2}$$

为描述简便，假设图像是正方形的，可使用张量积作为样条基，即

$$\boldsymbol{b}_{ij}(x,y)=\boldsymbol{b}_i(x)\otimes\boldsymbol{b}_j(y) \tag{4.3}$$

其中，$\boldsymbol{b}_i,\boldsymbol{b}_j$ 为与下述矩阵的较小的奇异值对应的奇异向量。

$$\boldsymbol{D}=\begin{bmatrix}1 & -2 & 1 & 0 & \cdots & 0\\ 0 & 1 & -2 & 1 & \cdots & 0\\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots\\ 0 & \cdots & \cdots & 1 & -2 & 1\end{bmatrix} \tag{4.4}$$

所谓张量积，也称外积，例如，

$$\boldsymbol{b}\otimes\boldsymbol{a}=\begin{bmatrix}b_1\\ b_2\\ b_3\\ b_4\end{bmatrix}\begin{bmatrix}a_1 & a_2 & a_3\end{bmatrix}=\begin{bmatrix}a_1b_1 & a_2b_1 & a_3b_1\\ a_1b_2 & a_2b_2 & a_3b_2\\ a_1b_3 & a_2b_3 & a_3b_3\\ a_1b_4 & a_2b_4 & a_3b_4\end{bmatrix} \tag{4.5}$$

对于平滑项，可以使用图像二次导数的平方和(能量项)。例如，对于基 $\boldsymbol{b}_{ij}(x,y)$ 的能量为

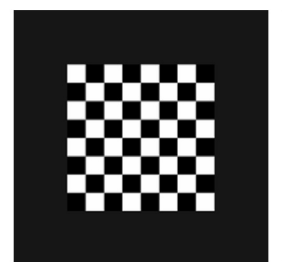
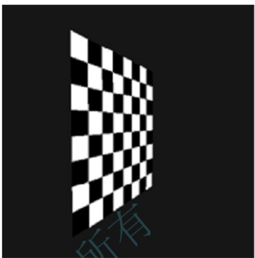
$$e_{ij}^2 = \|\boldsymbol{D}\boldsymbol{b}_i\|^2 + \|\boldsymbol{D}\boldsymbol{b}_j\|^2 \approx \iint \left(\frac{\mathrm{d}^2\boldsymbol{b}_{ij}}{\mathrm{d}x^2}\right)^2 + \left(\frac{\mathrm{d}^2\boldsymbol{b}_{ij}}{\mathrm{d}y^2}\right)^2 \mathrm{d}x\mathrm{d}y \tag{4.6}$$

其中，$\|\boldsymbol{b}_i\| = \|\boldsymbol{b}_j\| = 1$。

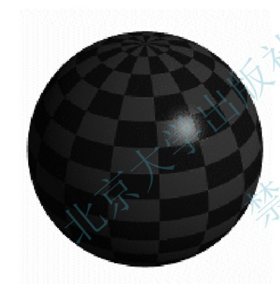
需要指出的是，对于张量积样条基，仅需要积分一维函数。向量 $\boldsymbol{D}\boldsymbol{b}_i$ 与 $\boldsymbol{D}$ 的左奇异向量成比例，因此是正交的。由于基函数 $\boldsymbol{b}_i$ 是正交的，就像 $\boldsymbol{D}\boldsymbol{b}_i$，由 $\boldsymbol{v}$ 控制的一个样条的平滑度可以写为 $\|\boldsymbol{E}\boldsymbol{v}\|^2$，其中 $\boldsymbol{E}$ 为一个对角矩阵，元素为每个基所用的 e_{ij}。这种方法相对于 Fourier 基的优势是，基向量是非循环的。这里所谓循环向量，假设 $\boldsymbol{V}$ 为有限维的向量空间，$\boldsymbol{A}$ 为一个自同态，若 $\boldsymbol{v}, \boldsymbol{A}\boldsymbol{v}, \cdots, \boldsymbol{A}^{n-1}\boldsymbol{v}$ 构成 $\boldsymbol{V}$ 的一个基，换句话说，pair$(\boldsymbol{A},\boldsymbol{v})$ 是完全可达的，则称 $\boldsymbol{v}$ 为 $\boldsymbol{V}$ 的一个循环向量。用一个相对低维的参数向量 $\boldsymbol{v}$ 来控制表面，可降低计算量，且可抑制含噪情况下的过拟合现象。

4.2.1 表面法向义性

使用传统的 SFT 方法估计表面法向时，主要根据投影过程发生的纹理畸变。例如，原来的正方形纹理经投影后可能变为平行四边形或退化的三角形，或者表面发生褶皱现象，如图 4.4 所示。

(a) 纹理元素形状畸变




(b) 纹理表面产生褶皱

图 4.4　纹理投影畸变

在正投影情况下，估计表面法向时通常存在双折叠倾斜歧义性，即法向分别为 $(p,q,1)$ 和 $(-p,-q,1)$ 的局部平坦面片的投影图像可能是相同的，如图 4.5 所示。

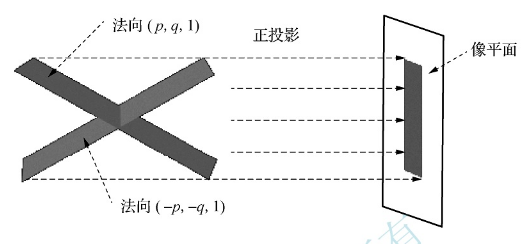
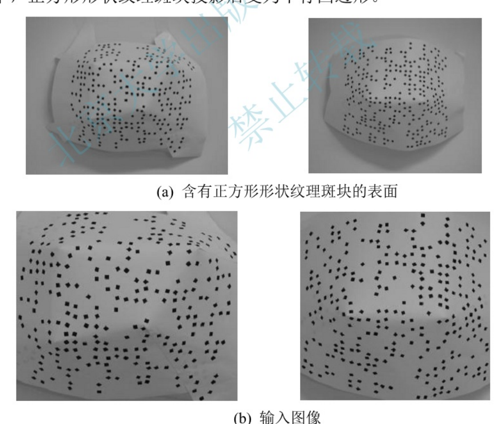


图 4.5　双折叠倾斜歧义性

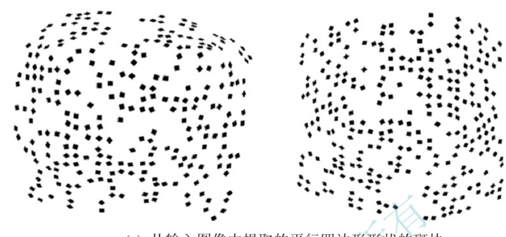
图 4.6 演示了具有大量歧义性法向表面的估计结果。这里使用简单的方法来恢复法向，即基于已知的纹理元素的畸变。例如，在图 4.6 中，正方形形状纹理斑块投影后变为平行四边形。



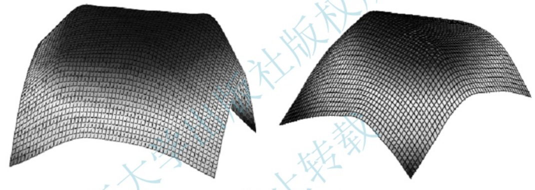
(a) 含有正方形形状纹理斑块的表面

(b) 输入图像

图 4.6　基于歧义性法向的表面形状估计



(c) 从输入图像中提取的平行四边形形状的斑块



(d) 计算得到的输出表面

图 4.6　基于歧义性法向的表面形状估计(续)

不过对于稀疏纹理情况，上述问题是严重约束不足的，因为对于法向的任何选择，表面都是连成一体的，而若存在纹理则将使得表面呈现多种视觉效果，看似存在各种形状。另外，在正投影下，存在一个全局的 in/out 倒置歧义性，以及绝对深度的连续歧义性，即深度信息为离散化。但如果表面是平滑的，可识别最可能的一种形状。另外，Forsyth[49]提出类似的交替算法，即在优化表面平滑度和选择法向之间交替执行。

若采用一个二次平滑项，上述问题将转化为一个彻底的离散优化问题。设表面的偏导数分别为 $p = \mathrm{d}z/\mathrm{d}x$ 、 $q = \mathrm{d}z/\mathrm{d}y$ 。在一个指定图

像点(x_i, y_i)观察到的纹理，为对应的表面导数提供了两种选择，记为$(p_i, q_i)d_i$，其中$d_i = \pm 1$，对应两个对折的表面。

将式(4.2)分别对x, y求偏导，可以用向量$\boldsymbol{v}$来表示样条表面的导数。

$$\begin{cases}(0,\cdots,0,p_i,0,\cdots,0)\boldsymbol{d} = \boldsymbol{a}_{p_i} \cdot \boldsymbol{v} \\ (0,\cdots,0,q_i,0,\cdots,0)\boldsymbol{d} = \boldsymbol{a}_{q_i} \cdot \boldsymbol{v}\end{cases} \tag{4.7}$$

其中，$\boldsymbol{d}$为一个由符号位(± 1) d_i构成的$n-$向量；而$\boldsymbol{a}_{p_i}, \boldsymbol{a}_{q_i}$为两个已知的行向量。

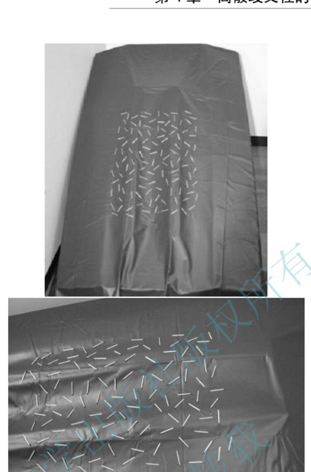
因此，使用一个二次平滑项，并通过使用矩阵$\boldsymbol{E}$将平滑项表示为样条参数的形式，这样，$\|\boldsymbol{E}\boldsymbol{v}\|^2$代表了平滑项能量。再用一个正则化参数$\lambda$来为平滑项能量指定权重，$\lambda$用来在平滑项能量和约束误差之间寻求均衡。总体来说，这些项可以写成类似式(4.1)的形式。

$$\arg\min_{\boldsymbol{v},\boldsymbol{d}} \left\| \underbrace{\begin{bmatrix} \sqrt{\lambda}\boldsymbol{E} \\ \boldsymbol{A}' \end{bmatrix}}_{\boldsymbol{A}} \boldsymbol{v} - \underbrace{\begin{bmatrix} 0 \\ \boldsymbol{B}' \end{bmatrix}}_{\boldsymbol{B}} \boldsymbol{d} \right\|^2 \tag{4.8}$$

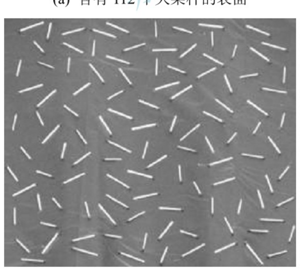
这里，$\boldsymbol{A}', \boldsymbol{B}'$为式(4.7)中的约束所构成的矩阵，即式(4.7)中的每个约束显示为$\boldsymbol{A}'\boldsymbol{v} = \boldsymbol{B}'\boldsymbol{d}$的一个单行。其中$\boldsymbol{v}$是连续的，$\boldsymbol{d}$是离散的。

4.2.2 线段定向歧义性

假设在一个平滑表面上检测出一组已知3D长度的线段，如图4.7所示。

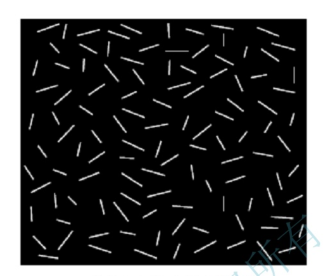


(a) 含有 112 个火柴杆的表面

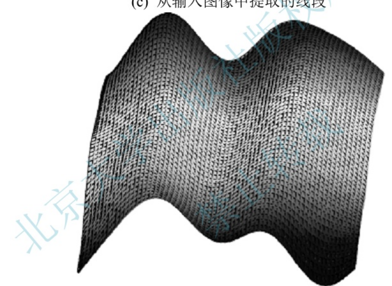


(b) 输入图像

图 4.7　线段及其表面计算结果



(c) 从输入图像中提取的线段



(d) 使用含 50 基函数的样条来计算得到的表面

图 4.7 线段及其表面计算结果(续)

给定一个正视图，每个线段具有一个 front/back 倒置，这就是定向歧义性，是离散化的。Naito 和 Rosenfeld[136]及 Koenderink 等人[137]研究了类似的问题。线段两个终点之间的深度差为

$$z_i - z_j = d_{ij}\sqrt{l^2 - r_{ij}^2} = \Delta_{ij} d_{ij} \tag{4.9}$$

其中，$d_{ij} = \pm 1$；l 是线段的 3D 长度；r_{ij} 是线段在图像中的长度。

假设l, r_{ij}(这里是Δ_{ij})是已知的。根据式(4.2)，点(x_i, y_i)处的深度z_i是该点样条基的线性组合。也就是说，z_i可表示为$\boldsymbol{a}_i \cdot \boldsymbol{v}$，其中$\boldsymbol{a}_i$为一个已知行向量。

每个深度约束可写为式(4.10)。

$$(\boldsymbol{a}_i - \boldsymbol{a}_j)\boldsymbol{v} = (0, \cdots, 0, \Delta_{ij}, 0, \cdots, 0)\boldsymbol{d} \tag{4.10}$$

收集所有约束的方程，加上平滑项，可以写成类似于式(4.8)的形式。

4.3 基于 SDP 的离散歧义性解决方案

下面求解式(4.1)形式的组合优化问题，如下式所示。

$$G: \quad \arg\min_{\boldsymbol{v},\boldsymbol{d}} \|A\boldsymbol{v} - B\boldsymbol{d}\|^2$$

其中，$\boldsymbol{A}, \boldsymbol{B}$是约束矩阵；$\boldsymbol{d} \in \{-1,1\}^n$是离散判定变量的一个向量，即$\boldsymbol{d} = (d_1, d_2, \cdots, d_n)$；$\boldsymbol{v} \in \mathbb{R}^m$是决定表面的连续变量的一个向量。

4.3.1 构造 SDP 松弛

下面设计一个 SDP，使得其值为式(4.1)的优化值 Opt(G) 的下限。这是最关键的一个步骤。若固定向量$\boldsymbol{d}$，则式(4.1)的优化解$\boldsymbol{v}$为

$$v = A^{+}Bd \tag{4.11}$$

其中，A^{+}为A的伪逆。再将式(4.11)反代入式(4.1)，得到

$$\begin{aligned}\left\|AA^{+}Bd - Bd\right\|^{2} &= \left\|(AA^{+}B - B)d\right\|^{2} \\ &= C \cdot X \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}\end{aligned} \tag{4.12}$$

其中，$C = (AA^{+}B - B)^{\mathrm{T}}(AA^{+}B - B) = B^{\mathrm{T}}(I - AA^{+})B$，$X = d^{\mathrm{T}}d$。因此，式(4.1)形式的问题归结为下面的仅关于$d$的优化问题。

$$\underset{d\in\{-1,1\}^{n}}{\arg\min}\, C \cdot (d^{\mathrm{T}}d) \tag{4.13}$$

式(4.13)与式(4.1)具有相同的优化值 Opt(G)。一旦求得了d，根据式(4.11)即可确定v，再根据式(4.2)即可得到表面 3D 形状。显然，如果$(v_1, v_2, \cdots, v_n, d_1, d_2, \cdots, d_n)$是式(4.1)的一个可行解，则$(d_1, d_2, \cdots, d_n)$为式(4.13)的一个可行解。

一般情况下，式(4.13)中的问题是一个 NP-难题[138]，无法在多项式时间内求的优化解。但是，很多文献从各个角度论证了 SDP 技术适用于求解此类问题的逼近解。

由于$d \in \{-1,1\}^{n}$，所以$d_i^2 = 1$，因此矩阵$X = d^{\mathrm{T}}d$是主对角线为 1 的对称半正定矩阵。这样，式(4.13)就可归结为下述 SDP 问题。

$$\underset{X}{\arg\min}\, C \cdot X \quad \text{s.t.} \quad x_{ii} = 1,\ i = 1, 2, \cdots, n, X \succeq 0 \tag{4.14}$$

换句话说，式(4.14)是式(4.13)的一种 SDP 松弛，且具有有限值 $\mathrm{SDP}(G) \leqslant \mathrm{Opt}(G)$。若对矩阵 $\boldsymbol{X}$ 进行 Cholesky 因式分解，即 $\boldsymbol{X}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$，其 $\boldsymbol{U}$ 为一个下三角矩阵。令 $\boldsymbol{u}_i$ 为 $\boldsymbol{U}$ 的第 i 列，由于 $x_{ij}=\boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{u}_j=1$，因此 $\boldsymbol{U}$ 中的列可看成将判定变量嵌入 $\mathbb{R}^n$ 中的一个单位球。

由于式(4.14)的 SDP 问题是凸性的，可以在多项式时间内用 SDP 解算机求解，可在多项式时间内找到一个 $x_{ii}^*=1$ 的矩阵 $\boldsymbol{X}^* \succeq 0$，使得

$$\boldsymbol{C}\cdot\boldsymbol{X}^* \leqslant \mathrm{SDP}(G)-\varepsilon \tag{4.15}$$

其中，ε 为任意小正数。

4.3.2　圆整操作

实际上我们希望求解式(4.13)，其中的 n 个变量 d_i 是 $S^0=\{-1,1\}$ 的元素，从而确定一个向量 $\boldsymbol{d}$。而现在拥有的是式(4.14)松弛规划的一个近似优化解 $\boldsymbol{X}^*$，前面讲过，通过 Cholesky 分解 $\boldsymbol{X}^*=\boldsymbol{U}^{*\mathrm{T}}\boldsymbol{U}^*$，得到的 $\boldsymbol{U}^*$ 含有 n 个向量变量 $\boldsymbol{u}_i \in S^{n-1}$。因此，需要寻找一种方法，将 S^{n-1} 反向映射到 S^0，在这样的过程中，要求不会丢失太多的目标函数值。

令 $\boldsymbol{p}\in S^{n-1}$，对应一个过原点的超平面，并定义下面的映射。

$$\begin{aligned}&\lambda_{\boldsymbol{p}}: S^{n-1}\to S^0\\ &\boldsymbol{u}\mapsto\begin{cases}+1 & ,\boldsymbol{p}^{\mathrm{T}}\boldsymbol{u}\geqslant 0\\ -1 & ,\boldsymbol{p}^{\mathrm{T}}\boldsymbol{u}<0\end{cases}\end{aligned} \tag{4.16}$$

λ_p 具有明确的几何解释，即 $\boldsymbol{p}$ 将 S^{n-1} 划分为一个闭合半球 $\boldsymbol{H}=\{\boldsymbol{u}\in S^{n-1}:\boldsymbol{p}^{\mathrm{T}}\boldsymbol{u}\geqslant 0\}$ 及其补半球 $\boldsymbol{H}^{-1}=\{\boldsymbol{u}\in S^{n-1}:\boldsymbol{p}^{\mathrm{T}}\boldsymbol{u}\leqslant 0\}$，如图4.8所示。

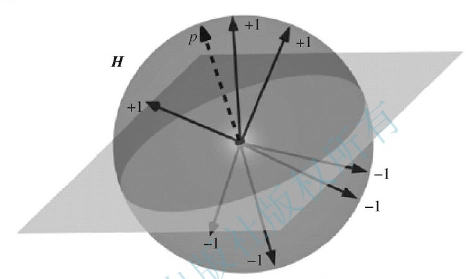


图4.8　圆整 S^{n-1} 中向量到 S^0

$\boldsymbol{H}$ 中的向量映射到1，而 $\boldsymbol{H}^{-1}$ 中的向量映射到−1。再随机选择一个 $\boldsymbol{p}$，重复执行上述过程。这个过程就称为圆整操作。

1. 随机法选择向量 $\boldsymbol{p}$

$\boldsymbol{p}$ 的更好的选取策略是，在 S^{n-1} 中随机均匀采样得到一个 $\boldsymbol{p}$。

为了不丢失太多的目标函数值，理所当然地希望 $\boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{u}_j$ 值较大的一对向量 $\boldsymbol{u}_i^*$ 和 $\boldsymbol{u}_j^*$ 比该值较小的一对向量更可能产生 x_{ij}。由于这个值随着 $\boldsymbol{u}_i^*$ 和 $\boldsymbol{u}_j^*$ 夹角的增大而增加，映射到 $\{-1,1\}$ 的措施将使得夹角大的那对向量将比夹角小的一对向量更可能映射到不同的值。

下面的定理将说明为什么随机选择 $\boldsymbol{p}$ 会使得式(4.16)具有较大的优势。

定理 4.1

令 $\boldsymbol{u},\boldsymbol{u}'\in S^{n-1}$，则式(4.16)将 $\boldsymbol{u},\boldsymbol{u}'$ 映射到不同值的概率为

$$\mathrm{prob}(\lambda_p(\boldsymbol{u})\neq\lambda_p(\boldsymbol{u}'))=\frac{1}{\pi}\arccos\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}' \tag{4.17}$$

证明：

令 $\alpha\in[0,\pi]$ 为单位向量 $\boldsymbol{u}$ 和 $\boldsymbol{u}'$ 的夹角，根据余弦定律，有

$$\cos(\alpha)=\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}'\in[-1,1] \tag{4.18}$$

换句话说，$\alpha=\arccos(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}')\in[0,\pi]$。

如果 $\alpha=0$ 或 $\alpha=\pi$，意味着 $\boldsymbol{u}\in\{\boldsymbol{u}',-\boldsymbol{u}'\}$，这个结论较易满足。否则，考察 $\boldsymbol{u}$ 和 $\boldsymbol{u}'$ 的一种线性翼展 L，即 $L=\mathrm{span}(\boldsymbol{u},\boldsymbol{u}')$，这是 $\mathbb{R}^n$ 的一个二维子空间。

令 $\boldsymbol{r}$ 为 $\boldsymbol{p}$ 在该子空间的投影，则有

$$\begin{cases}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{u}=\boldsymbol{r}^{\mathrm{T}}\boldsymbol{u}\\ \boldsymbol{p}^{\mathrm{T}}\boldsymbol{u}'=\boldsymbol{r}^{\mathrm{T}}\boldsymbol{u}'\end{cases} \tag{4.19}$$

这意味着，$\lambda_p(\boldsymbol{u})\neq\lambda_p(\boldsymbol{u}')$，当且仅当 $\boldsymbol{r}$ 位于开度角 α 的一个半开双面楔子 W 中，即图 4.9 所示的阴影区域。

由于 $\boldsymbol{p}$ 在 S^{n-1} 中均匀分布，$\boldsymbol{r}$ 的方向在 $[0,2\pi]$ 中也呈均匀分布。

因此，$\boldsymbol{r}$ 落进双楔子的概率为双楔子所覆盖的角范围所占的比率，即 α/π。

□

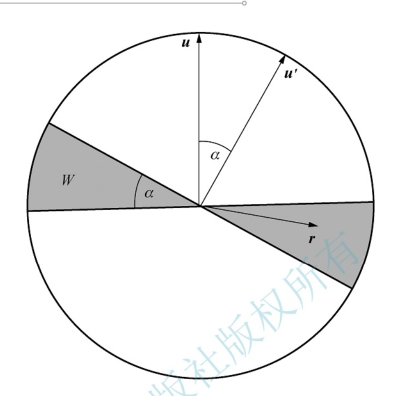


图 4.9　随机圆整

通过上述圆整步骤，即可从连续解矩阵 $\boldsymbol{X}$ 中得到离散向量 $\boldsymbol{d}$ 。

2. 分裂惯性法选择向量 $\boldsymbol{p}$

下面提出一种新的方案用于选择向量 $\boldsymbol{p}\in S^{n-1}$ ，即通过分裂惯性法来分割单位球。所谓分裂惯性法(Splitting Inertial Methods，SIM)，是由回归坐标对分(Recursive Coordinate Bisection，RCB)[139]加工而成，用于空间图形分割。在 3D 空间，使用下述矩阵的最小特征值所对应的特征向量作为分割轴。

$$\boldsymbol{I}=\begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \tag{4.20}$$

其中，

$$I_{xx}=\sum_{v\in N}((y(v)-y_c)^2+(z(v)-z_c)^2), I_{yy}=\sum_{v\in N}((x(v)-x_c)^2+(z(v)-z_c)^2)$$

$$I_{zz}=\sum_{v\in N}((x(v)-x_c)^2+(y(v)-y_c)^2), I_{xy}=I_{yx}=-\sum_{v\in N}(x(v)-x_c)(y(v)-y_c)$$

$$I_{yz}=I_{zy}=-\sum_{v\in N}(y(v)-y_c)(z(v)-z_c), I_{xz}=I_{zx}=-\sum_{v\in N}(x(v)-x_c)(z(v)-z_c)$$

$$(x_c,y_c,z_c)=\frac{\sum_{v\in N}(x(v),y(v),z(v))}{|N|}$$

上面式子中的 $(x(v),y(v),z(v))$ 表示点 v 的 3D 坐标，N 为顶点的总数目。

需要找到一个平面，垂直于最小角动量的轴，将图中的顶点划分为两个相等的子集。这种方法和 K-L[143]方法成功结合，即在每个循环步骤中，用分裂惯性法得到的结果再用 K-L 法加以提炼。分裂惯性法以前也被用于其他的嵌入操作[141,142]，但还没有用于 SDP 嵌入的。

结合本书的研究对象，设 $\boldsymbol{X}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$，使用 $\boldsymbol{U}$ 的相当于最大奇异值的主奇异向量(Principle Singular Vectors，PSV)。事实上，如果必须选择一个平面的话，一个对法向有益的选择就是那个主奇异向量。

为了扩大平面的选择面，选取相当于 k 个最大奇异值的奇异向量的一个线性加权组合作为平面法向选取的法向，即

$$\boldsymbol{p}=\sum s_i\lambda_i\boldsymbol{w}_i \tag{4.21}$$

其中，$\boldsymbol{w}_i$ 是相当于奇异值 λ_i 的奇异向量。而 $s_i \sim \boldsymbol{N}(0,1)$，表示可随机取值。

下面简单说明如何求解上述奇异向量。在大多数线性代数中，都是通过求解特征多项式的根来得到特征值。可在大多数情况下，这种特征多项式的根对多项式系数的变化很敏感，所以需要另辟蹊径。

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，$\boldsymbol{D}$ 为一个对角矩阵，如果存在一个可逆矩阵 $\boldsymbol{X}$，满足 $\boldsymbol{A} = \boldsymbol{XDX}^{-1}$，则称 $\boldsymbol{A}$ 是可对角化的。如果 $\boldsymbol{A}$ 是可对角化的，当且仅当 $\boldsymbol{A}$ 含有 n 个线性无关特征向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$，构成 $\mathbb{R}^n$ 的一个基。因此，令

$$\boldsymbol{v}^{(0)} = a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_n\boldsymbol{x}_n \tag{4.22}$$

其中的 $a_1, a_2, \cdots, a_n$ 是标量。在上式两边同时乘以 $\boldsymbol{A}^k$，得到

$$\begin{aligned} A^k\boldsymbol{v}^{(0)} &= A^k(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_n\boldsymbol{x}_n) \\ &= a_1\lambda_1^k\boldsymbol{x}_1 + a_2\lambda_2^k\boldsymbol{x}_2 + \cdots + a_n\lambda_n^k\boldsymbol{x}_n = a_1\lambda_1^k\left(\boldsymbol{x}_1 + \sum_{j=2}^{n}\frac{a_j}{a_1}\left(\frac{\lambda_j}{\lambda_1}\right)^k\boldsymbol{x}_j\right) \end{aligned} \tag{4.23}$$

如果 $\boldsymbol{A}$ 含有有序特征值 $\lambda_1, \cdots, \lambda_n$，即 $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$，则 λ_1 称为主特征值。显然，有

$$\lim_{k\to\infty}\left(\frac{\lambda_j}{\lambda_1}\right)^k = 0 \tag{4.24}$$

因此，如果 $a_1 \neq 0$，则 $\boldsymbol{A}^k\boldsymbol{v}^{(0)} \to a_1\lambda_1^k\boldsymbol{x}_1$。

可见，上式将正则化，避免了上溢或下溢，因此，收敛。

将上述分析过程总结为表 4-2 所示的迭代算法。

表 4-2　求解最大特征值、最大特征向量的迭代算法

步骤	内容
1	令 $\boldsymbol{v}^{(0)}$ 为单位向量
2	For　$k=1,2,\cdots$ (1)　$\boldsymbol{w}=\boldsymbol{A}\boldsymbol{v}^{(k-1)}$ (2)　$\boldsymbol{v}^{(k)}=\boldsymbol{w}/\\|\boldsymbol{w}\\|$ (3)　$\lambda^{(k)}=(\boldsymbol{v}^{(k)})^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}^{(k)}$
3	返回结果

由于$|\lambda_1|\geqslant|\lambda_2|\geqslant\cdots\geqslant|\lambda_n|$，因此，$\lim\limits_{k\to\infty}(\lambda_j/\lambda_1)^k\to 0$。又由于特征向量$\boldsymbol{v}^{(k)}\propto\boldsymbol{e}_1$，所以当$k\to\infty$，$\boldsymbol{v}^{(k)}$收敛于一个主向量。通过上述迭代方法，可以有效求解特征向量，进而获得上述奇异向量。

3. *角扫描*

下面在上述离散向量的判定策略上再进行改进，即不是根据单个向量$\boldsymbol{p}$来判定，而根据式(4.21)来随机选择一对法向$\boldsymbol{p}_1,\boldsymbol{p}_2$，所有由$\boldsymbol{p}\in\mathrm{Span}(\boldsymbol{p}_1,\boldsymbol{p}_2)$产生的法向$\boldsymbol{p}\in\mathbb{R}^2$构成一个单位圆，即$\|\boldsymbol{p}\|_2=1$。该单位圆过图 4.8 的球心，将球分成两个半球，形成一个平面。将嵌入$\mathbb{R}^n$的点投影到这个平面。

使用极坐标系，用$\boldsymbol{\theta}=(\theta_1,\theta_2,\cdots,\theta_n)^{\mathrm{T}}\in\mathbb{R}^n$表示$\mathbb{R}^2$中的$n$个向量$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$，即

$$\boldsymbol{p}_i=\begin{pmatrix}\cos\theta_i\\ \sin\theta_i\end{pmatrix}，\forall i=1,2,\cdots,n \tag{4.25}$$

显然，有

$$\boldsymbol{p}_i^{\mathrm{T}}\boldsymbol{p}_j \equiv \cos(\theta_i-\theta_j) \quad ,\forall i,j=1,2,\cdots,n \tag{4.26}$$

令 $T(\theta)$ 为斜对称矩阵值函数，即

$$T_{ij}(\theta)=\theta_i-\theta_j \tag{4.27}$$

再定义映射 $f:\mathbb{R}^n\to\mathbb{R}$ 。

$$f(\theta)\equiv\frac{1}{2}\boldsymbol{W}\cdot\cos(T(\theta)) \quad ,\forall\theta\in\mathbb{R}^n \tag{4.28}$$

其中， $\cos(T(\theta))$ 为 $n\times n$ 矩阵，其元素为 $T(\theta)$ 的元素的余弦。

这样，式(4.14)的问题就转化为下述形式。

$$\underset{\theta\in\mathbb{R}^n}{\arg\min}\, f(\theta) \tag{4.29}$$

这是一个无约束问题，一般具有多重局部、非全局极小值。

很容易计算 $f(\theta)$ 的导数。 $f(\theta)$ 的一阶偏导数为

$$\frac{\partial f(\theta)}{\partial\theta_j}=\sum_{k=1}^{n}w_{kj}\sin(\theta_k-\theta_j) \quad ,\forall j=1,2,\cdots,n \tag{4.30}$$

因此，有

$$g(\theta)\equiv\nabla f(\theta)=\left[\boldsymbol{W}\circ\sin(T(\theta))\right]^{\mathrm{T}}\mathrm{e} \tag{4.31}$$

其中，“∘”表示 Hadamard，即 $\sin(T(\theta))$ 和 $\boldsymbol{W}$ 逐元素相乘。

$f(\theta)$ 的二阶偏导数为

$$\frac{\partial^2 f(\theta)}{\partial\theta_i\partial\theta_j}=\begin{cases} w_{ij}\cos(\theta_i-\theta_j) & ,i\neq j \\ -\sum\limits_{k\neq j}w_{kj}\cos(\theta_k-\theta_j) & ,i=j \end{cases} \tag{4.32}$$

因此，$f(\theta)$ 的 Hessian 为

$$H(\theta) \equiv \nabla^2 f(\theta) = \boldsymbol{W} \circ \cos(T(\theta)) - \mathrm{diag}([\boldsymbol{W} \circ \cos(T(\theta))]\mathrm{e}) \tag{4.33}$$

其中，对于任一个向量 $\boldsymbol{p}$，$\mathrm{diag}(\boldsymbol{p})$ 是指对角为 $\boldsymbol{p}$ 的对角矩阵。

令 $\bar{\theta} \in \mathbb{R}^n$ 为向量之间的一种角表示，即存在一个整数 k_{ij}，满足下式。

$$\bar{\theta}_i - \bar{\theta}_j = k_{ij}\pi \tag{4.34}$$

很显然，$\cos(\bar{\theta}_i - \bar{\theta}_j) = \pm 1$，并且存在一个二进制向量 $\bar{\boldsymbol{x}} \in \{-1,1\}^n$，满足

$$\cos(\bar{\theta}_i - \bar{\theta}_j) \equiv \bar{x}_i \bar{x}_j = \pm 1$$

当围绕 θ 的每个分量均匀旋转，$f(\theta)$ 将保持不变，即对任何标量 τ，$f(\theta) \equiv f(\theta + \tau\mathrm{e})$。同时，$f(\theta)$ 对于每个 θ_i 呈周期性，周期为 2π。

若为均匀旋转和周期性建模，发现在二进制和角表示之间存在明显的一一映射关系。

$$\bar{\theta}_i = \begin{cases} 0 & , \bar{x}_i = +1 \\ \pi & , \bar{x}_i = -1 \end{cases} \tag{4.35}$$

由于对于任何满足式(4.34)的 $\bar{\theta}$ 有 $\sin(\bar{\theta}_i - \bar{\theta}_j) = 0$，可直接从式(4.31)得出 $g(\bar{\theta}) = 0$，即 $\bar{\theta}$ 为 $f(\theta)$ 的一个驻点。

这样，需要解决的问题就转换为最大图割问题。为了得到近似优化解，首先极小化 $f(\theta)$，得到一个局部极小值 θ，相当于点在单位圆上的一种分布。利用 $f(\theta)$ 的周期性，假设 $\theta_i \in [0, 2\pi]$，且 $\theta_1 \leqslant \theta_2 \leqslant \cdots \leqslant \theta_n$。

将单位圆划分的两个相等部分的任何一个划分都对应一个图割。例如，选取一个角 $\alpha \in [0,\pi]$，则可根据下式进行划分。

$$x_i = \begin{cases} +1 & ,\theta_i \in [\alpha, \alpha+\pi] \\ -1 & ,\theta_i \notin [\alpha, \alpha+\pi] \end{cases} \tag{4.36}$$

上述最大割算法总结如表 4-3 所示。设当前图割 $\boldsymbol{x}$ 的值为 γ，输入 θ，输出 x^*。

表 4-3　最大割算法

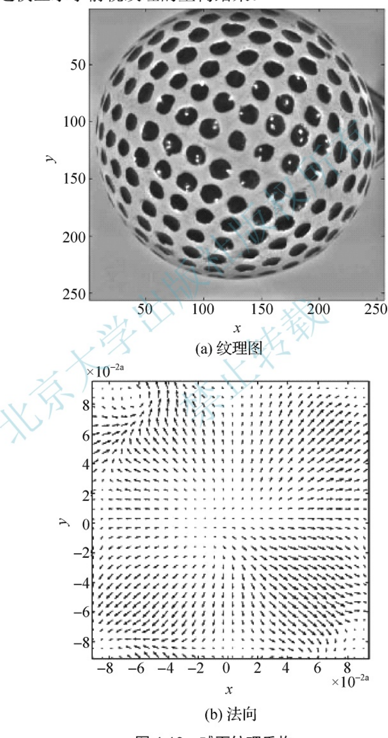
步骤	内容
1	令 $\alpha = 0, \Gamma = -\infty, i = 1$
2	令 j 为最小的序号，$\theta_j > \pi$，$j = j+1$，$\theta_{n+1} = 2\pi$
3	While $\alpha \leqslant \pi$ (1) 根据式(4.36)计算 $\boldsymbol{x}$，同时计算 $\gamma(\boldsymbol{x})$。 (2) 若 $\gamma(\boldsymbol{x}) > \Gamma$，则 $\Gamma = \gamma(\boldsymbol{x})$，$\boldsymbol{x}^* = \boldsymbol{x}$。 (3) 若 $\theta_i \leqslant \theta_j - \pi$，则 $\alpha = \theta_i, i = i+1$；否则，$\alpha = \theta_j - \pi, j = j+1$
4	返回结果

上述角扫描方法比前述球面法更高效。基于单个随机法向分裂 $\mathbb{R}^n$ 中 n 个点的成本是 $O(n^2)$。如果这些点投影到一个平面，并且一个最初割已被计算出来，则可以将割成本降低为 $O(n)$。通过类似于 K-L 算法[143]的提炼过程，可获得更精确的估计结果。使用 K-L 算法进行改良是一个局部搜索过程，可以消除少量放错位置的顶点。如果在 50 次连续迭代中没有进展，则将提早结束算法[144]。最后，返回最低成本的解。

图 4.6 描述了从具有双折叠歧义性法向的纹理畸变实现表面重构。为简化纹理提取过程，使用正方形结构的纹理元素。

图 4.10 描述了球面问题的重构。其中，图 4.10(a)的球面含有 89

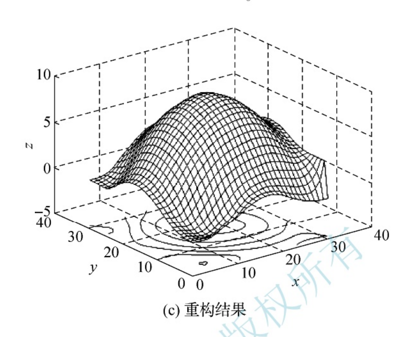
个纹理元素(不包含遮挡超过 50%的纹理元素)；图 4.10(b)显示了对应的法向，一定程度上体现了球面纹理的统计特性；图 4.10(c)是重构结果，不过仅显示了前视纹理的重构结果。



(a) 纹理图

(b) 法向

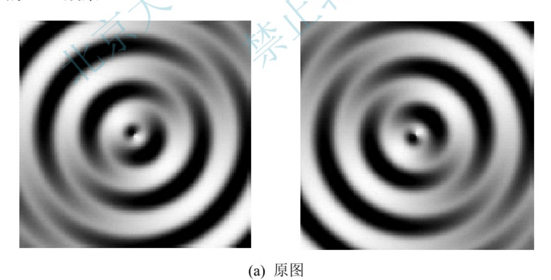
图 4.10　球面纹理重构



(c) 重构结果

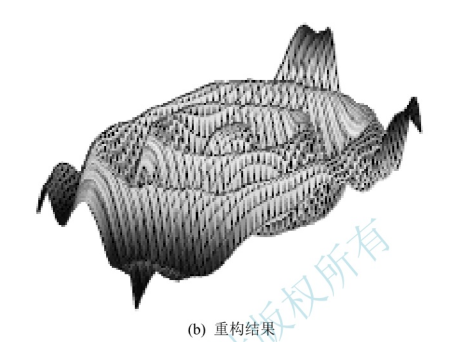
图 4.10 球面纹理重构(续)

图 4.11 演示了从一个 Lambertian 表面的一对合成图像计算出一个表面。原图为立体光测度所得的结果，重构结果反映了双折叠歧义性的 SFT 效果。



(a) 原图

图 4.11 从一个 Lambertian 表面的一对合成图像计算出一个表面



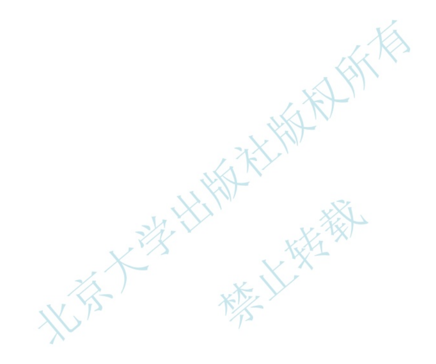
(b) 重构结果

图 4.11　从一个 Lambertian 表面的一对合成图像计算出一个表面(续)

4.4　本 章 小 结

由于 SFT 方法是严重约束不足问题，一般的假定约束只能用于某些特定领域或解决某个特定问题。尤其当表面纹理图案存在严重歧义时，常规 SFT 方法将导致多重歧义解，对解的合理选取及表面积分都造成极大的计算负担。本章主要研究了二进制歧义性问题，以双折叠歧义性为例，提出将由纹理畸变线索判断表面平滑性的双折叠歧义性问题归结为一个离散判定问题，将表面表示为一个张量积基函数的线性组合，再用离散变量描述该歧义性约束，进而定义一个二次目标函数来度量表面平滑度和约束满足度，通过消除连续变量产生一个仅含离散变量的二次目标函数，再使用 SDP 凸松弛法，将离散变量嵌入一

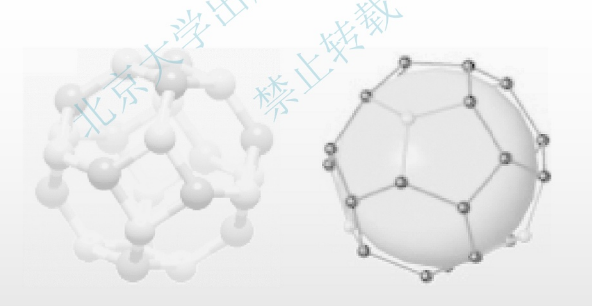
个连续高维空间，通过基于分裂惯性和角扫描的圆整化算法将离散变量归约到一个低维空间，最后通过最大割算法获得较理想的解。上述方案比传统的椭球法高效且成本低($O(n^2)$)，可进一步使用 K-L 算法改良局部搜索过程，消除错误割点，有效提高解的精确度。





第5章

基于共平面线性约束的表面重构



共平面性是 2D 图像中几何元素的一种重要特征，包括平行性和共线性。基于共平面性的 3D 重构不同于前面介绍的基于纹理畸变(SFT)和明暗渐变性(SFS)的重构策略，属于纯粹的几何学范畴，常用于图像分析与识别。本章着重介绍其在 3D 表面重构中的应用原理及具体的实现技术。

5.1 引　　言

基于单图像的 3D 表面重构由于其严重的病态性，需要附加特殊的假定条件。例如，假设表面曲线是平坦的，或者位于多面体的平坦削面，或者位于多面体的某个平坦横截面，如图 5.1 和图 5.2 所示。

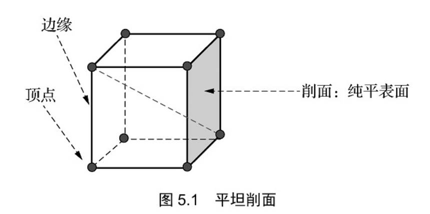


图 5.1　平坦削面

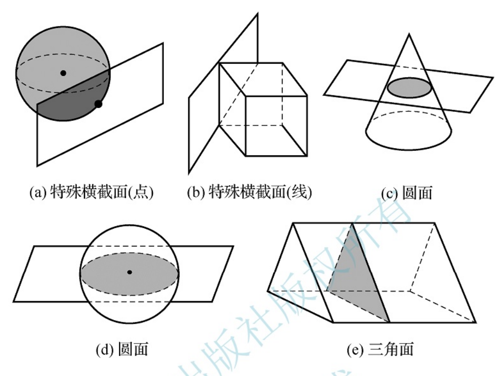
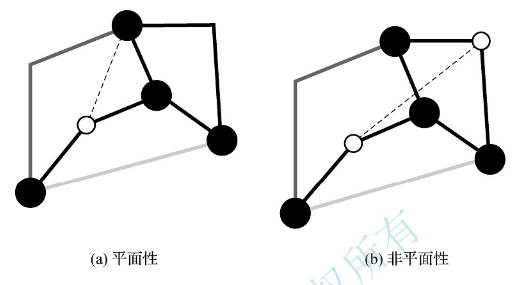


图 5.2　几何横截面

共平面性由于其特殊的几何性质被广泛应用于距离遥测、摄影测量、心理学、辅助医疗等领域。下面先简单回顾一下平面性的基本原理及基本的约束模式。

5.1.1　平面性

所谓图形的平面性，是指从任何角度观察，该图形都可以被绘制在一个平面上，且图形边缘除了顶点之外没有其他的交叉点，如图 5.3 所示。



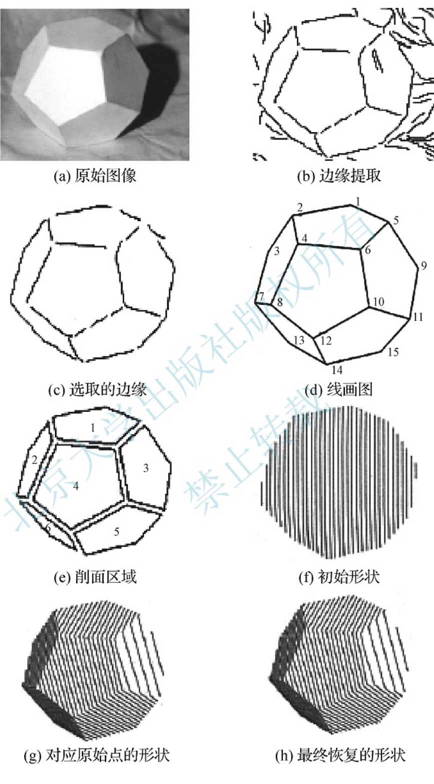
(a) 平面性　　(b) 非平面性

图 5.3　图形的平面性与非平面性

对于人类视觉感知的研究表明[145,146]，人类视觉系统通常使用平面性假设，平面性甚至可以替代立体视觉的深度信息用于感知三维场景[147]。事实上，真实环境中的平坦曲线十分普遍，尤其在人造环境中更是普遍存在，如建筑物、道路等。但是，平坦曲线在自然界中却特别少见，不过，地平面或平静的水面等也可看成是平坦的。另外，空间物体的运动轨迹也大都呈现近似平坦的。

在计算机视觉领域，最早使用平面性约束的是 Roberts[56]，首次将这种特殊的几何性质用于基于图像的 3D 表面重构研究中。图 5.4 描述了基于平面性的表面重构的基本原理。

从图 5.4 可见，Roberts 主要针对已知形状的多面体对象。由于多面体中平坦削面的交叉线在投影图像上生成直边缘，相对易于检测，从而将 3D 表面的重构问题归结为从线画图构造一种 3D 描述。



(a) 原始图像　(b) 边缘提取

(c) 选取的边缘　(d) 线画图

(e) 削面区域　(f) 初始形状

(g) 对应原始点的形状　(h) 最终恢复的形状

图 5.4　基于线画图的多面体表面重构

之后 Guzman[148]试图解决未知对象的识别问题，为此定义了顶点的不同类型，如图 5.5 所示。

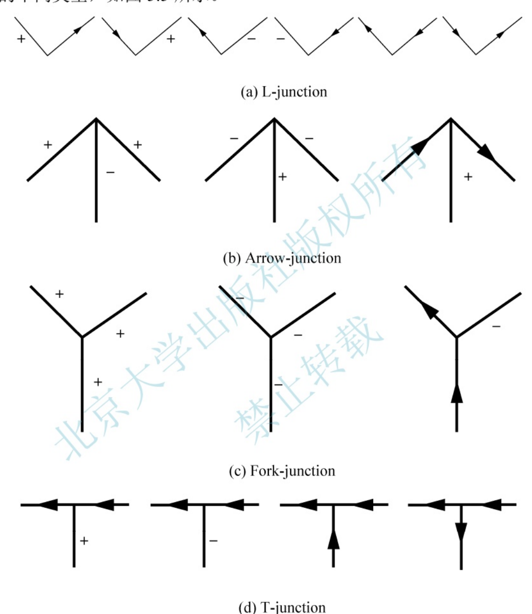


图 5.5　Guzman 定义的顶点类型

在图 5.5 中，“+”表示凸边缘，“-”表示凹边缘，实际效果如图 5.6 所示。

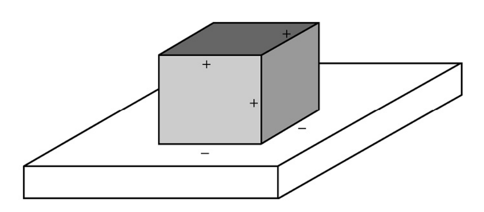


图 5.6　凸边缘与凹边缘

如图 5.7 所示，L-junction 和 Arrow-junction 类型的顶点是指对象中遮挡了一个背景面的拐角点；而 Fork-junction 类型没有遮挡背景面；一个对象的被遮挡的轮廓与被遮挡的对象的轮廓形成 T-junction 类型的顶点。

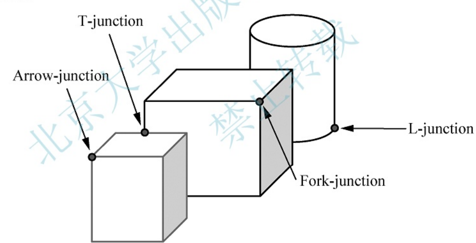


图 5.7　不同类型的顶点示例

Guzman 指出，如果所有边缘属于单个 3D 物体对象，那么，某种类型的顶点不能仅由某一条边缘链接，并用 Penrose 多边形等不可能的光幻觉图形进行了测试，如图 5.8 所示。

(a) 三角形

(b) 正方形

(c) 五边形

(d) 六边形

(e) 八边形

图 5.8 Penrose 多边形

Huffman[149]和 Clowes[150]分别独自研究了三面顶点的某些更荒谬的组合构型，如图 5.9 所示。

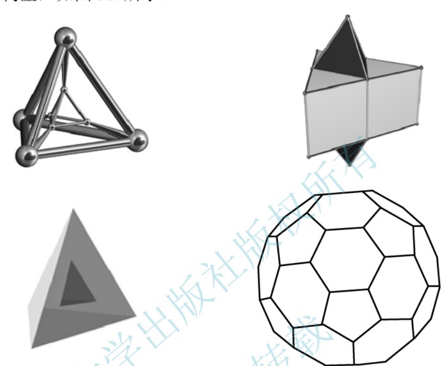


图 5.9　三面顶点的特殊构型

他们采用所谓的 Huffman-Clowes 标签，将每个边缘标记为凸、凹或遮挡等三种形态，如图 5.10 所示。

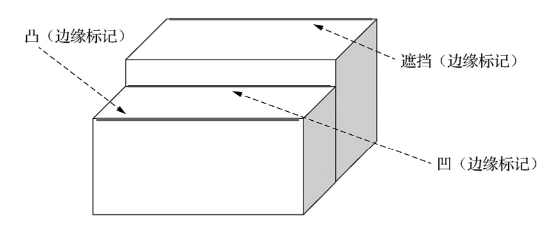


图 5.10　Huffman-Clowes 标签

这样一来，一个顶点链表就包含了交汇于该有效三面顶点的边缘的标签，从而将判断一个线画图能否被 3D 识别的问题转化为对每个边缘分配一个标签。尽管标签问题是一个经典的 NP-难题[151]，但根据 Waltz 的研究[152]，围绕一个顶点分配标签后，邻域顶点的数目大大减少了，就可使用其他的优化方法加以逼近求解。

Malik[153]将上述线标签模式扩展到曲面研究，而 Varley[154]将顶点链表扩展到四面顶点，如图 5.11 所示。

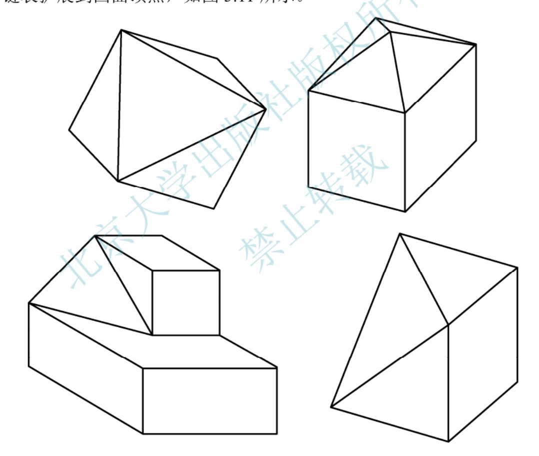


图 5.11　四面顶点的特殊构型

直线标签存在尚无定论的理论及技术问题。首先，根据直线标签不能直接获得 3D 表面；其次，四面顶点的理想标签的平均数目对于顶点数目呈指数级增长[155]，因此直线标签不足以描述 3D 形状；第三，

一个理想标签是 3D 重构的必须但非充分的条件，因为存在一些虚假的线画图，不能识别为有效的 3D 对象。

5.1.2　共平面性

由于直线标签的线画图重构具有许多未完全解决的理论和技术问题，Sugihara[156]提出在此基础上增加代数结构的共平面性约束条件。所谓 3D 空间中的一个点集是共平面的，是指所有这些点位于同一个几何平面。例如，3 个点必定是共平面的，而 4 个以上的点可能位于不同的平面，则是非共平面的。

同样，3D 空间中的两条直线如果平行或交叉，则也是共平面的，如图 5.12 所示。

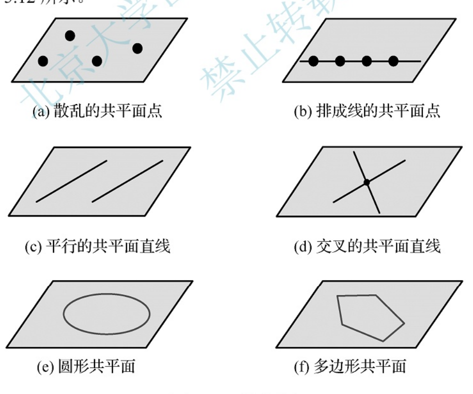
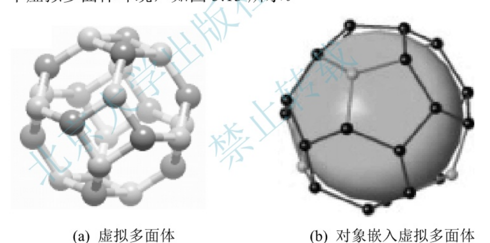


图 5.12　共平面性

Sugihara 同时指出，这是将一个线画图识别为一个 3D 多面体的充要条件，而 Falk[157]实际构造了一个共平面性约束的线性系统，但未从数学上根本解决问题。

共平面性蕴含了点的共线性和线的平行性，如图 5.12(b)、图 5.12(c)所示。一般地，图像中的平行线既可能是 3D 平行线的投影(在正投影下)，也可能是共平面 3D 线的投影(透视投影下)。同样，图像中的近似共线点可以假设为 3D 空间的共面点。之所以强调这一点，是因为自然界中的共线性与平行线情况比平坦表面更为普遍。

Rothwell 等人[158,159,160]提出多面体笼子的比喻，即将一个 3D 对象用一个虚拟多面体环绕，如图 5.13 所示。



(a) 虚拟多面体　　(b) 对象嵌入虚拟多面体

图 5.13　多面体笼子

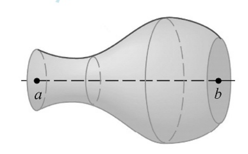
这个虚拟多面体的顶点不需要贴附在所环绕对象的表面上，但要确保可从该对象表面的图像中计算出虚拟多面体的顶点坐标。由于位于多面体顶点的点集的投影呈视角不变性(称为视常量)，因此一旦虚拟多面体的顶点被识别出来，就可以计算出一个视常量，再通过查询一个预设的形状数据库轻易地识别出对象的类别[161]。

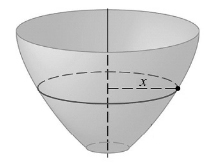
但是，Burns 等人研究指出[162]，并非对于任意大小的点集都存在一个视常量。不过，这并不意味着对于某些特殊的点集构型也不存在视角无关。存在视常量的条件研究得最多的是共平面性[159,163]。

在正投影情况下，对于 4 个共平面点存在视常量；而对于透视投影，5 个共平面点存在视常量。事实上，这些视常量仅对 4 个以上的共平面点才有意义，因为三个点是自然的共平面性，而一个“多面体笼子”必然存在部分至少含有 4 个顶点的削面。在这样的削面上，可以定义仿射不变量。而在两个交叉平面上，可以定义透视不变量。

5.1.3　平坦横截面

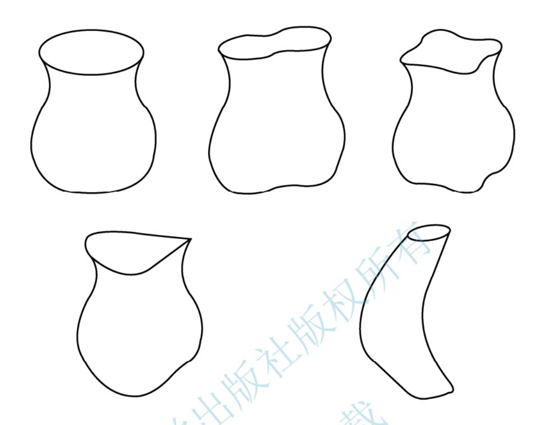
另一种用于分析多面体的是平坦横截面。具有平坦横截面的对象很普遍，如回转曲面(Surfaces Of Revolution，SOR)[164]、直同构广义柱面(Straight Homogeneous Generalized Cylinders，SHGC)[165]，如图 5.14 所示。





(a) 回转曲面，由一条曲线绕一个轴线圆周旋转而成

图 5.14　平坦横截面示例



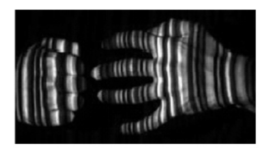
(b) 直同构广义柱面，有一条直线轴，横截面形状固定，但尺寸大小可变

图 5.14 平坦横截面示例(续)

基于激光扫描的分析系统就是典型的横截面分析例子。图 5.15 描述了完全未标定结构光的成像及重构效果。

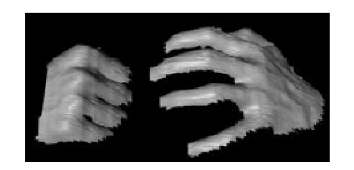


(a) 3D 手型



(b) 激光扫描

图 5.15 基于未标定结构光的 3D 重构



(c) 计算得到的 3D 表面

图 5.15　基于未标定结构光的 3D 重构(续)

在图 5.15 中，通过若干个样条激光扫描手腕，相当于做横切操作，得到一系列横截面，其中的图像曲线就是横截面的投影。根据横截面图像可计算出 3D 表面，而不需要平面的其他任何约束信息。

5.1.4　共平面约束

共平面性关系可用于计算机视觉领域的很多实际应用中，迄今已涌现出很多基于共平面约束的分析模式。在无噪声情况下，这些模式在数学上是等价的，但在计算代价上却是各有千秋。

最早由 Sugihara[156]提出的约束模式包含了所有的自由变量，即

$$\boldsymbol{v}=(a_1,\cdots,a_N,b_1,\cdots,b_N,d_1,\cdots,d_N,z_1,\cdots,z_k) \tag{5.1}$$

在结构光情形下，可能有 $k=O(N^2)$ 个交叉点，但仅需要 $3N$ 个参数来表示平面。

Lipson 等人[166]通过平面拟合来构造共平面约束，得到 m 个共平面点的深度 $\boldsymbol{z}=(z_1,\cdots,z_m)^{\mathrm{T}}$ 上的线性约束。由于对于共平面点，存在 a,b,d，满足

$$\begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_m & y_m & 1 \end{bmatrix}\begin{pmatrix} a \\ b \\ d \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} = \boldsymbol{z} \tag{5.2}$$

若令

$$\boldsymbol{P} = \begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_m & y_m & 1 \end{bmatrix} \tag{5.3}$$

则有

$$\boldsymbol{P}^{+}\boldsymbol{z} = \begin{pmatrix} a \\ b \\ d \end{pmatrix} \tag{5.4}$$

$$(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^{+})\boldsymbol{z} = 0 \tag{5.5}$$

其中，$\boldsymbol{P}^{+}$ 是 $\boldsymbol{P}$ 的伪逆。如果构造 $\boldsymbol{S}_i = (\boldsymbol{I} - \boldsymbol{P}_i\boldsymbol{P}_i^{+})$，其中 $\boldsymbol{P}_i$ 包含了每个平坦剖面的点。再用 0 填满这些矩阵，则下面的矩阵

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{s}_1 \\ \vdots \\ \boldsymbol{s}_N \end{bmatrix}$$

描述了所有共平面性约束，即 $\boldsymbol{S}\boldsymbol{z} = 0$，其中 $\boldsymbol{z} = (z_1, \cdots, z_k)^{\mathrm{T}}$。在这种形式下，不仅矩阵 $\boldsymbol{S}$ 的大小为 $O(k \times k)$，而且较稠密。

Sparr[167]和 Heyden[168]等人相继提出共面点深度的稀疏约束，使用质心表示。对于每个平坦剖面，选择 3 个非共线点 (x_1, y_1)、(x_2, y_2)、(x_3, y_3) 作为一个基。该剖面上图像坐标为 (x_4, y_4) 的任何其他点可表示为式(5.6)。

$$\begin{pmatrix} x_4 \\ y_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \alpha\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} - \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) + \beta\left(\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} - \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) \tag{5.6}$$

经重新整理，得到

$$c_1\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + c_2\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + c_3\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} + c_4\begin{pmatrix} x_4 \\ y_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{5.7}$$

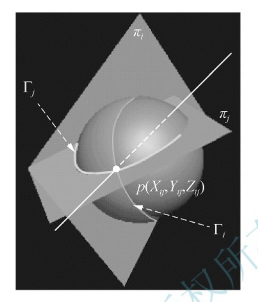
其中，$c_1 + c_2 + c_3 + c_4 = 0$。

上述质心对于仿射变换保持恒定不变。平面上所有点的约束收集在一起构成矩阵 $\boldsymbol{S}$，满足 $\boldsymbol{S}z = 0$。

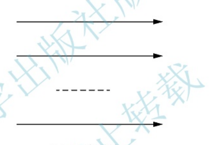
5.2　基于共平面性的线性约束系统

假设一个表面 $\boldsymbol{F}$ 上含有 N 个平坦曲线 $\Gamma_1,\cdots,\Gamma_N$，分别嵌入 N 个平面 $\pi_1,\cdots,\pi_N$ 上。每条曲线 Γ_i 含有平面 π_i 上的一个点集，换句话说，曲线 Γ_i 是一组共平面点。现在假设在正投影下得到一幅图像，从两条曲线 Γ_i,Γ_j 在像平面上的投影图像上识别出 k 个交叉点 (x_{ij}, y_{ij})，如图 5.16 所示。

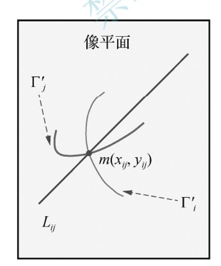
如果能识别出图像点 (x_{ij}, y_{ij}) 进而求出交叉曲线 Γ_i,Γ_j，就可以恢复对应的平面 π_i,π_j。需要说明的是，两条曲线 Γ_i,Γ_j 可能在图像中不相交，也可能交叉多次。如果两条曲线共享一个直线段，则将直线段的两个端点作为交叉点。当两条以上的曲线交于一点，需要按一种次序轮流考察其中的每一对曲线。



(a) 平面与表面相交产生曲线



(b) 正投影



(c) 曲线的正投影图像

图 5.16　3D 空间交叉平坦曲线及其投影

5.2.1　构造线性系统

当平面 π_i 不平行于投影方向，则可表示为一个线性方程。

$$z_i(x,y)=a_i x+b_i y+d_i \tag{5.8}$$

其中 z_i 相当于平面 π_i 上某点 (x_i,y_i) 处的深度信息。在正投影下，图像中的一个 2D 交叉点 $m(x_{ij},y_{ij})$，相当于表面上的一个 3D 交叉点，其对应的 X,Y 坐标等效于 (x_{ij},y_{ij})，这样该 3D 点在对应的两个平面 π_i,π_j 上的 Z 轴方向的坐标之差即为该 3D 点的深度差，即

$$z_i(x_{ij},y_{ij})-z_j(x_{ij},yi_j)=(a_i-a_j)x_{ij}+(b_i-b_j)y_{ij}+(d_i-d_j) \tag{5.9}$$

由于该 3D 点是两个平面与表面相交而形成的曲线的交点，所以有

$$z_i(x_{ij},y_{ij})-z_j(x_{ij},y_{ij})=0 \tag{5.10}$$

实质上，平面 π_i,π_j 在空间交叉于一条 3D 直线，该相交线在像平面上的投影也为直线，定义为 L_{ij}，满足下述方程。

$$(a_i-a_j)x+(b_i-b_j)y+(d_i-d_j)=0 \tag{5.11}$$

因此，根据式(5.9)、式(5.10)和式(5.11)，意味着投影点 (x_{ij},y_{ij}) 位于直线 L_{ij} 上。

考察所有可见的交叉点，将对应的式(5.10)方程组成一个线性方程组，即

$$\boldsymbol{A}\boldsymbol{v}=\boldsymbol{0} \tag{5.12}$$

其中，向量$\boldsymbol{v}=(a_1,\cdots,a_N,b_1,\cdots,b_N,d_1,\cdots,d_N)^{\mathrm{T}}$为相交平面簇的参数向量，一个$\boldsymbol{v}$定义了表面曲线的集合$\{\Gamma_i\}$；$\boldsymbol{A}$是一个由 5 个元素$x_{ij},-x_{ij},y_{ij},-y_{ij},1,-1$等构成的$M\times N$的稀疏矩阵，$M$表示曲线交点的数目。

类似地，对于透视投影情况，具有类似的线性系统结构。若设焦距为f，图像坐标x,y与世界坐标X,Y之间的关系为$x=fX/Z,y=fY/Z$。设平面π_i不穿越相机中心，则有

$$a_iX+b_iY+c_iZ(X,Y)=\left(\frac{a_ix}{f}+\frac{b_iy}{f}+c_i\right)Z(X,Y)=1 \tag{5.13}$$

再根据一对曲线的相交关系，产生如下类似于式(5.10)的线性方程。

$$\frac{1}{Z_i(x_{ij},y_{ij})}-\frac{1}{Z_j(x_{ij},y_{ij})}=\frac{(a_i-a_j)x_{ij}}{f}+\frac{(b_i-b_j)y_{ij}}{f}+(c_i-c_j)=0 \tag{5.14}$$

对所有可见的交叉点，组合多个式(5.14)方程，构成如下类似于式(5.12)的线性方程。

$$\boldsymbol{A}_f\boldsymbol{v}_f=\boldsymbol{0} \tag{5.15}$$

其中，$\boldsymbol{A}_f$由将$\boldsymbol{A}$矩阵的元素x_{ij},y_{ij}除以f而得到。

5.2.2　平凡解子空间

根据前面的分析，由于 $\boldsymbol{v}=(a_1,\cdots,a_N,b_1,\cdots,b_N,d_1,\cdots,d_N)^{\mathrm{T}}$，因此系统最多含有 $3N$ 个变量。正投影情况下，一组真实平面构成的向量 $\boldsymbol{v}$ 必然是式(5.12)的一个解，并且存在于 Null($\boldsymbol{A}$) 中。需要注意的是，Null($\boldsymbol{A}$) 也可能包含某些平凡解。例如，当所有曲线同时被嵌入同一个平面时，就产生一个平凡解。另外，通过观察可见，式(5.10)具有一个基本平凡解子空间，表示为 $\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$，其中，

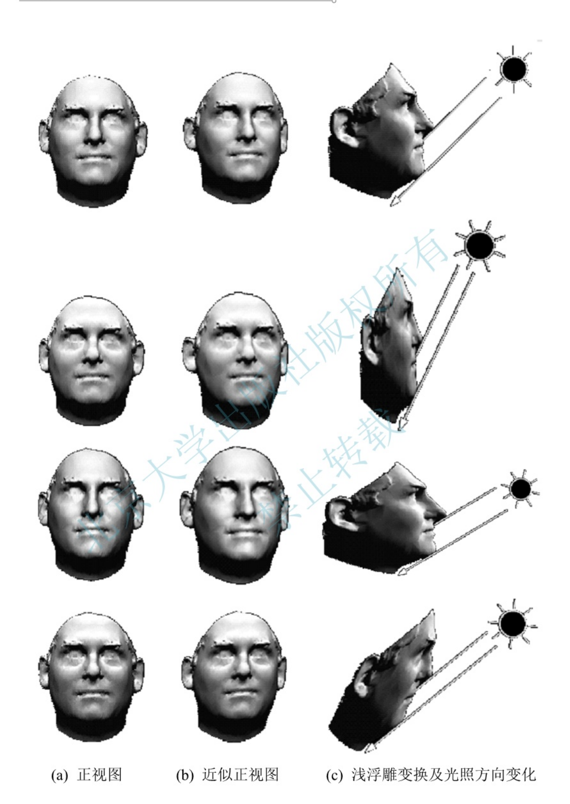
$$\boldsymbol{v}_1=\frac{(\mathbf{1}_N,\mathbf{0}_{2N})^{\mathrm{T}}}{\sqrt{N}},\boldsymbol{v}_2=\frac{(\mathbf{0}_N,\mathbf{1}_N,\mathbf{0}_N)^{\mathrm{T}}}{\sqrt{N}},\boldsymbol{v}_3=\frac{(\mathbf{0}_{2N},\mathbf{1}_N)^{\mathrm{T}}}{\sqrt{N}} \tag{5.16}$$

式(5.16)中，$\mathbf{1}_N$、$\mathbf{0}_N$、$\mathbf{0}_{2N}$ 分别表示元素值为 1、0 的 $N-$向量和 0 的 $2N-$向量。

这样，对于式(5.12)的任一个非纯平解 $\boldsymbol{v}$，存在歧义解，即 $\boldsymbol{v},\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 的任一个线性组合 $\boldsymbol{t}=\alpha_1\boldsymbol{v}+\alpha_2\boldsymbol{v}_1+\alpha_3\boldsymbol{v}_2+\alpha_4\boldsymbol{v}_3$ 也是式(5.12)的一个解。这就是广义浅浮雕(Generalized Bas-Relief，GBR)歧义性[169]。图 5.17 显示了人脸浅浮雕[170]的歧义性实例。

由图 5.17 可见，即使改变了光照方向和浮雕参数，得到的正视图效果几乎无差异。

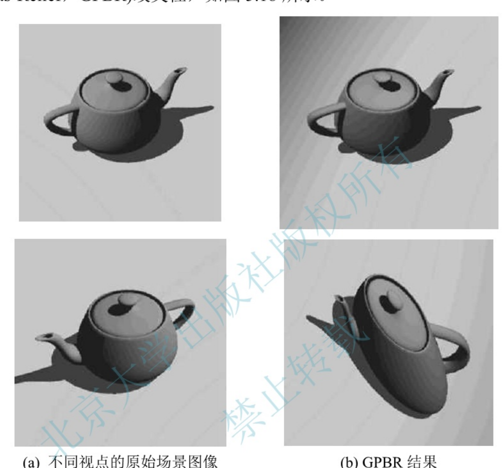
GBR 给人的直观感觉是，为表面上所有点增加一个平面，或者将深度乘以一个常量，以维持共平面关系，但观察者看不见。



(a) 正视图　　(b) 近似正视图　　(c) 浅浮雕变换及光照方向变化

图 5.17　GBR 歧义性

在透视投影下，存在广义透视浅浮雕(Generalized Perspective Bas-Relief，GPBR)歧义性，如图 5.18 所示。



(a) 不同视点的原始场景图像 (b) GPBR 结果

图 5.18 GPBR 歧义性

浅浮雕歧义性在一些未标定深度的环境中很普通，如明暗感知、光度测定立体，以及基于两视图的仿射重构[171]等。Koenderink 指出[50]，不同的观察者对同一张图片的 3D 感知是不同的，其原因就是正投影的浅浮雕歧义性。

从投影几何变换角度来说，GBR 和 GPBR 是 3D 空间一种特殊的仿射和射影变换。GBR 变换可以写为式(5.17)。

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ d \end{pmatrix} \tag{5.17}$$

而 GPBR 可以写成式(5.18)的齐次坐标变换形式。

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ W' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{5.18}$$

为了理解 GPBR 如何作用于平面，考察平面 $a_1X + b_1Y + c_1Z = 1$ 上的点 $P = (X_0, Y_0, Z_0)^{\mathrm{T}}$，该点的齐次坐标为 $(X_0, Y_0, Z_0, 1)^{\mathrm{T}} = (X_0, Y_0, Z_0, a_1X_0 + b_1Y_0 + c_1Z_0)^{\mathrm{T}}$。根据式(5.18)，得到

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ W' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix} \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \\ a_1X_0 + b_1Y_0 + c_1Z_0 \end{pmatrix}$$

$$= \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \\ (a + a_1d)X_0 + (b + b_1d)Y_0 + (c + c_1d)Z_0 \end{pmatrix} \tag{5.19}$$

如果 $W' \neq 0$，则点 (X', Y', Z', W') 在下面的平面上。

$$(a + a_1d)X + (b + b_1d)Y + (c + c_1d)Z = 1 \tag{5.20}$$

这样，在 GPBR 变换下，式(5.12)的某个解可表示为式(5.21)的线性组合。

$$t = a\sqrt{N}\boldsymbol{v}_1 + b\sqrt{N}\boldsymbol{v}_2 + c\sqrt{N}\boldsymbol{v}_3 + d\boldsymbol{v} \tag{5.21}$$

类似的研究工作还有很多，如 Sugihara 等[156]主要研究多面物体的平坦削面，而有文献[165]指出，基于平坦横截面的线性系统主要用于结构光和直同构广义柱面。需要指出的是，一个 SHGC 的平坦横截面可不必在图像中可见。某些情况下，横截面可根据轮廓线估计出来[172]。

5.3 无噪系统的非平凡解

在 5.2 节讨论了基于共平面性约束而构成的线性系统的平凡解子空间。下面，再考察该系统的非平凡解子空间。在无噪的理想情况下，式(5.12)存在一个四维解空间，即含有 4 个线性无关解，其中的 3 维 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 源自于 GBR 歧义性，而第 4 维 $\boldsymbol{v}$ 来自于真实的形状。根据前面的分析，式(5.12)解空间的维数依赖于观察到的曲线交叉的样式，部分不相交的曲线导致了更多的自由度。

定理 5.1(自由度下限)

设 D 为一个标签化的线画图，S 是由 D 表示的一个空间结构。如果 S 与 D 一起描述了一个多面体，且 S 具有两个以上的削面，那么，根据 S 和 D 来确定的一个多面体至少具有 4 个以上的自由度。

证明：

令 P 为构成由 S 和 D 表示的多面体的 3D 点集，P' 为通过下述放射变换得到的所有的点集。

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & \beta & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \delta \end{bmatrix} \tag{5.22}$$

其中，α,β,δ 是任一个实数，而 γ 是正实数，点 (x,y,z) 是 P 上的要变换的点。由于是仿射变换，P 中的边缘和削面分别变换为 P' 的边缘和削面，因此，P' 也是一个多面体。P' 和 P 一样拥有 S 和 D，因为，仿射变换不改变原始点的 x,y 坐标，并且式(5.22)的系数矩阵的行列式是正的(因为 $\gamma>0$)，所以，该变换不改变空间位置，也就是说，削面的外法向从不会倒置。这样一来，P' 也是由 S 和 D 表示的一个多面体。

式(5.22)的变换具有 4 个参数，如果 P 上的 4 个非共面点的对应变换点能被确定，则又唯一确定这 4 个参数值。由于多面体 P 含有两个以上的削面，总是可以选取 4 个非共面点。因此，如果希望从 S 和 D 重建多面体，可以任意选择至少 4 个顶点的 z 值。

□

Rothwell 等人[158]验证了含有 4 个以上自由度但交叉较少的多面体情况。特殊的情况是，如果图像是一个三角形网格，可以将顶点放置到任意深度中。但是，对于高连通曲线网络，系统将表现为典型的超定，即方程组中方程的个数大于未知变量的数量，无法完全满足给定条件，此时只能得到一个最接近解。例如，在一个结构光情境中，得到 $O(N^2)$ 方程，含有 $3N$ 个未知数。尽管极端罕见，但仍然存在一种高连通系统，具有 $O(N^2)$ 个交叉，具有一个非平凡解的子空间，子空间的维数大于 1。

下面使用几何方法来人工构造一个含有 $O(N^2)$ 个交叉点，但具有一个 2D 非平凡解子空间的系统。构造过程如下。

令 $\{\pi_i^1\}$、$\{\pi_i^2\}$ 为两个含有 N 个平面的集合。对于每个平面集合，将每一对平面交叉，并寻找交叉线在图像中的投影［参见式(5.11)］。每个平面集合的这些交叉线的投影图像线的集合分别定义为 $\{L_{ij}^1\}$、$\{L_{ij}^2\}$。设某一对 L_{ij}^1, L_{ij}^2 的交叉点为 (x_{ij}, y_{ij})。通过上述构造过程，显然平面集合 $\{\pi_i^1\}$、$\{\pi_i^2\}$ 就是这些交叉点 (x_{ij}, y_{ij}) 所形成的方程组的解。不过具有三维非平凡子空间的系统更加罕见，因为每组交叉线 $L_{ij}^1, L_{ij}^2, L_{ij}^3$ 必须交于一点，或者两条线重合。

5.4　含噪系统的非平凡解

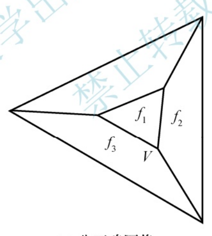
实际应用中噪声是普遍存在的。Sugihara[156]指出，在含噪情况下，线性系统是典型的超定，没有精确的非平凡解，只有逼近解。处理这类超严格问题，有两种方法。

第一种方法是假设空间曲线是严格平坦(称为纯平)的，所以，交叉点在图像中的投影点的位置的测量值一般需要被矫正，即测量值不可能精确，均需要矫正到一个对应的纯平面内。假设一个已标签化的线画图 D，$I=(V,F,R)$ 是其对应的关联结构，其中 V,F 为互不相交的有限点集，R 为序偶集合 $V\times F$ 的一个子集，表示一个弧线集合。如果 $(v,f)\in R$，表示“顶点 v 位于削面 f 之上”或“削面 f 含有顶点 v”。令 R^* 为 R 的最大子集。表 5-1 描述了顶点的矫正算法，输入：D，$I=(V,F,R)$。输出：矫正的线画图。

表 5-1　顶点位置矫正算法

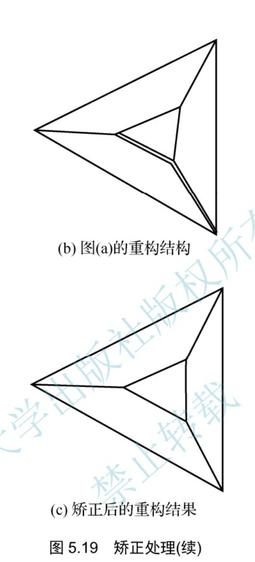
步骤	内容
1	寻找 R 的最大子集 R^*，构造 $I=(V,F,R^*)$
2	求解类似式(5.12)的 $\boldsymbol{A}_{R^*}\boldsymbol{v}=0$，如果无解，返回 FALSE
3	寻找 $V(R-R^*)$ 中的顶点的正确的空间位置，作为表面的交叉点
4	将这些矫正后的空间位置投影到像平面
5	观察矫正后的图像是否服从既定标签规则，若服从，返回 TRUE；否则返回 FALSE
6	采用矫正点重新构成式(5.12)，若 $\boldsymbol{v}^*$ 满足新的方程系统，返回校正后的图像；否则返回 FALSE

图 5.19 是一个矫正处理的实例。



(a) 非正确图像

图 5.19　矫正处理



(b) 图(a)的重构结构

(c) 矫正后的重构结果

图 5.19　矫正处理(续)

目前的很多方法都属于这类范畴。Sugihara 等人通过寻找方程组的一个子集，精简这些方程结构，确保即使交点存在微小的干扰，系统仍然有解。但是，Sugihara 的方法不能均衡地传播误差。Shimshoni 等人[174]进而提出线性规划松弛法，提供了一个用于曲线非平凡解识别的必要但非充分的条件。而 Ros 等人[173]则研究了非线性优化算法，基于对所有精确解的一种特殊的参数化方法，即根据顶点和削面的一种排序，寻求一种途径来在 3D 空间依次放置这些顶点和削面。

Heyden[168]首次提出一种针对超严格问题的数学方法，即对方程组的任一个不相容子集，使用 SVD 来修改系统矩阵。

第二种方法是从完全相反的角度来考虑，即假设交点位置度量值是完全正确的，也就是说式(5.12)中的 $\boldsymbol{A}$ 是固定的，而作为一种折中处理，反过来假设曲线是非严格平坦的。这种假设也许更符合自然界的实际情况。这种方法的明显优势是，可采用数值线性代数方法。相比之下，矫正线画图的方法组合度更高，而且通常是一个 NP-难题。目前沿着这个方向开展研究的有 Grimstead[175]、Sturm[25]、Bouguet[176]等。

5.4.1 平凡子空间

平凡子空间通常指零空间和全空间自身。基于式(5.16)中的向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$，可能并不能生成纯平解的整个子空间，也就是说，可能还有其他的平凡解。例如，当某条曲线与其他曲线的交叉点排列在一条直线上，特别是对于交叉点小于 3 个的一条曲线，将发生这种情况，因为此时 $\mathrm{Span}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)$ 不具普遍性，可能遗漏解空间中的某些向量，如图 5.20 所示。

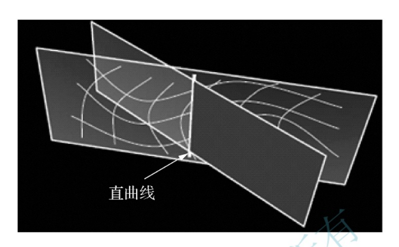
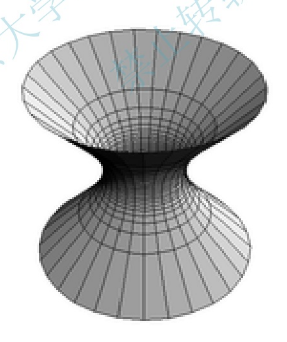


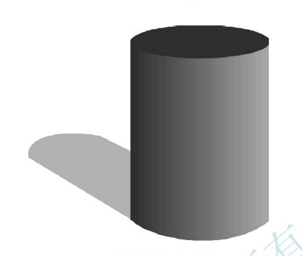
图 5.20　直曲线

如果嵌入了直曲线的那个平面不平行于投影方向，则该直曲线对应一条 3D 直线。但需要说明的是，零高斯曲率(Zero Gaussian Curvature，ZGC)表面[175]是一种特殊的表面，其中也包含了直线，不属于上述情况，如图 5.21 所示。

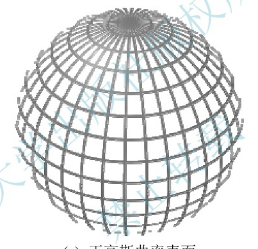


(a) 负高斯曲率表面

图 5.21　不同高斯曲率的表面实例



(b) 零高斯曲率表面



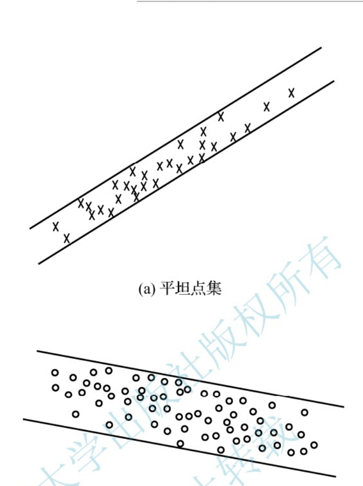
(c) 正高斯曲率表面

图 5.21 不同高斯曲率的表面实例(续)

1. 度量表面平坦度

为了描述纯平解的完备空间，需要度量表面的平坦度。有很多方法可用于估计表面的平坦度，最常用的是电子水准仪、激光干涉仪、自动瞄准仪等。

在数学上，若考察 N 数据点的平坦度，可用两个平行平面将数据点集夹在中间，求解这两个平面之间的距离的最小值，如图 5.22 所示。



(a) 平坦点集

(b) 非平坦点集

图 5.22　平坦度示例

比较简单的方法是，采用线性回归法根据采集的有代表性的数据点拟合出一个平面，并度量其残差。如果所有点共面，则残差为 0。将采集到的点的深度描述为 $\boldsymbol{v}$ 的一个线性函数。令 (x_i, y_i) 为曲线上 k 个点的集合。再令 $\boldsymbol{Z}$ 为一个矩阵，用于反投影这些点到它们的对应平面。$\boldsymbol{Z}$ 的行在适当的列中包含 $x_i, y_i, 1$，则有 $\boldsymbol{Z}\boldsymbol{v} = (z_1, \cdots, z_k)^{\mathrm{T}}$，其中 $(z_1, \cdots, z_k)^{\mathrm{T}}$ 就是平面在点 (x_i, y_i) 处的深度信息。

定义矩阵 $\boldsymbol{C}$ 为

$$\boldsymbol{C}=\frac{(\boldsymbol{I}-\boldsymbol{P}\boldsymbol{P}^{+})\boldsymbol{Z}}{\sqrt{k}}=\frac{\boldsymbol{Z}-\boldsymbol{P}(\boldsymbol{P}^{+}\boldsymbol{Z})}{\sqrt{k}} \tag{5.23}$$

其中，

$$\boldsymbol{P}=\begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_k & y_k & 1 \end{bmatrix}$$

$\boldsymbol{P}^{+}$ 表示 $\boldsymbol{P}$ 的伪逆。那么，$\|\boldsymbol{C}\boldsymbol{v}\|^2$ 为点 (x_i,y_i,z_i) 在 Z 方向偏离最佳拟合平面的偏差的范数的均方值。这种线性回归的残差等价于求解下述的最小二乘问题。

$$\min\frac{1}{k}\left\|\boldsymbol{Z}\boldsymbol{v}-\boldsymbol{P}\begin{pmatrix} a \\ b \\ d \end{pmatrix}\right\|^2 \tag{5.24}$$

若令 $(a,b,d)^{\mathrm{T}}=\boldsymbol{P}^{+}\boldsymbol{Z}\boldsymbol{v}$，则式(5.24)的残差为

$$\frac{1}{k}\|(\boldsymbol{Z}-\boldsymbol{P}\boldsymbol{P}^{+}\boldsymbol{Z})\boldsymbol{v}\|^2=\|\boldsymbol{C}\boldsymbol{v}\|^2 \tag{5.25}$$

2. 平凡子空间

定义平凡子空间为 $\boldsymbol{C}$ 的零空间，即 $\mathrm{Null}(\boldsymbol{C})=\{\boldsymbol{v}\in\boldsymbol{V}:\boldsymbol{C}\boldsymbol{v}=\boldsymbol{0}\}$。平凡子空间的所有向量相当于空间的共面点集 (x_i,y_i,z_i)，即平凡解。由于非共线的三点确定一个平面，所以选择这个共面存在 3 个自由度，但是，$\mathrm{Null}(\boldsymbol{C})$ 的维数可能比较大。

这个平凡子空间的选择具有较大的自由度，但可由某些附加信息来加以限制，如某些已知点的深度。假设一组随机平面被从一个随机方向正交观察，或者这些平面具有随机的均匀的法向，那么这些基本平凡分量是如何分布的呢？前面讲过，式(5.26)中的$\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$是基本平凡子空间中的一组标准正交基向量。基本平凡分量的相对度量为

$$\frac{\left\|(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}_1, \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}_2, \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}_3)\right\|}{\left\|\boldsymbol{v}\right\|} \tag{5.26}$$

由于在正交视图中丢失了绝对深度，可令

$$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}_3 = \frac{1}{\sqrt{N}}\sum d_i = 0 \tag{5.27}$$

则式(5.26)的上限为

$$\frac{\left\|\left(\sum a_i, \sum b_i\right)\right\|}{\sqrt{N}\left\|(a_1,\cdots,a_N,b_1,\cdots,b_N)\right\|} \tag{5.28}$$

当平面方向均匀分布时，未标度的法向的法向为$(n_i^x, n_i^y, -n_i^z)$，其中$n_i^x, n_i^y, n_i^z \sim \mathrm{N}(0,1)$，即服从期望为 0、方差为 1 的正态分布。

在式(5.8)形式下，有$a_i \sim n_i^x / n_i^z$、$b_i \sim n_i^y / n_i^z$。平面的系数呈 Cauchy 分布。Cauchy 分布是一种连续稳态分布，不存在均值、方差和高阶矩。因此，不能将中心极限定理用于平凡分量$\frac{1}{N}\sum a_i$或$\frac{1}{N}\sum a_i$的绝对度量，并期望当$N \to \infty$时，趋向于正态分布。但是，式(5.26)的基本平凡分量的相对度量的界限总为 1。

显然，若存在大量的法向均匀分布的平面，$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}_1$和$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}_2$的度量可能比$\boldsymbol{v}$的度量要小。因此，如果没有附加信息，在正投影的情形下，

可以选择一个正交于该平凡子空间的解 $\boldsymbol{v}$，即该 $\boldsymbol{v}$ 与平凡子空间的每个向量都正交，记为 $\boldsymbol{v} \perp \text{Null}(\boldsymbol{C})$，得到一个高质量的解。

5.4.2 正投影系统的求解

在含噪情况下，平凡解仍然是精确的。但在一般情况下，没有另外的精确解，即只能得到近似解。从几何学上看，这意味着在点 (x_{ij}, y_{ij}) 处平面之间存在深度间隙。

解决的基本方法是，极小化式(5.12)中的残差的范数，即寻找一个向量 $\boldsymbol{v}$，使得 $\|\boldsymbol{A}\boldsymbol{v}\|$ 较小。由于当空间数据点 (x_i, y_i, z_i) 是位于一个近似平坦面内时，$\|\boldsymbol{A}\boldsymbol{v}\|$ 对任何曲线集都是较小的，无法模式化求解。因此，要求点 (x_i, y_i, z_i) 远离一个公共平面。根据有关文献[92]，$\|\boldsymbol{C}\boldsymbol{v}\| = 1$ 是使点集远离其最佳拟合平面，即可能的公共平面的充分条件。

$\boldsymbol{C}$ 的零空间 $\text{Null}(\boldsymbol{C})$ 由平凡解构成，而这些平凡解可拟合为一个单平面，即平凡子空间中的所有向量相当于共面点集。再根据式(5.23)、式(5.24)，$\boldsymbol{C}\boldsymbol{v} = 0$ 意味着 $\boldsymbol{Z}\boldsymbol{v} = \boldsymbol{P}(a, b, d)^{\mathrm{T}}$，其中 a, b, d 为平面参数，即 $z_i(x_{ij}, y_{ij}) = z_j(x_{ij}, y_{ij}) = ax_{ij} + by_{ij} + d$。根据定义，这意味着 $\boldsymbol{A}\boldsymbol{v} = 0$。即

$$\boldsymbol{C}\boldsymbol{v} = 0 \;\Rightarrow\; \boldsymbol{A}\boldsymbol{v} = 0 \tag{5.29}$$

因此，$\text{Null}(\boldsymbol{C}) \subseteq \text{Null}(\boldsymbol{A})$。

现在，根据第 5.4.1 节的分析结论，可以假定 $\boldsymbol{v}$ 正交于 $\text{Null}(\boldsymbol{C})$，因为平凡子空间的任何解 $\boldsymbol{v}$ 不影响 $\|\boldsymbol{A}\boldsymbol{v}\|$。至此，可将式(5.12)改写为如式(5.30)。

$$\arg\min_{v} \|\boldsymbol{Av}\| \quad \text{s.t.} \quad \|\boldsymbol{Cv}\| = 1 \text{ , } \boldsymbol{v} \perp \text{Null}(\boldsymbol{C}) \tag{5.30}$$

为了求解式(5.30)，令 $\boldsymbol{U}\tilde{\boldsymbol{D}}\tilde{\boldsymbol{V}}^{\mathrm{T}}$ 为 $\boldsymbol{C}$ 的奇异值分解。通过分别移除相当于小于 ε 的奇异值的 $\tilde{\boldsymbol{V}}$ 中的列和 $\tilde{\boldsymbol{D}}$ 中的行，得到 $\boldsymbol{V},\boldsymbol{D}$。$\tilde{\boldsymbol{V}}$ 中被移除的列翼展后覆盖了平凡子空间，该平凡子空间至少是三维。$\boldsymbol{V}$ 的列恰好构成平凡子空间的正交补的一个基。

若令 $\boldsymbol{v} = \boldsymbol{VD}^{-1}\boldsymbol{w}$，由于 $\boldsymbol{U}$ 是正交的，这样，上述问题变换为

$$\arg\min_{w} \|\boldsymbol{AVD}^{-1}\boldsymbol{w}\| \quad \text{s.t.} \quad \|\boldsymbol{UDV}^{\mathrm{T}}\boldsymbol{v}\| = \|\boldsymbol{w}\| = 1 \tag{5.31}$$

上述正投影系统的求解算法总结如表 5-2 所示。

表 5-2　正投影系统的求解算法

步骤	内容
1	根据交叉点集，构造矩阵 $\boldsymbol{A},\boldsymbol{Z},\boldsymbol{P}$
2	令 $\boldsymbol{C} = [\boldsymbol{Z} - \boldsymbol{P}(\boldsymbol{P}^{+}\boldsymbol{Z})]/\sqrt{k}$
3	计算 $\boldsymbol{C}$ 的 SVD：　$\boldsymbol{C} = \boldsymbol{U}\tilde{\boldsymbol{D}}\tilde{\boldsymbol{V}}^{\mathrm{T}}$
4	根据 $\tilde{\boldsymbol{D}},\tilde{\boldsymbol{V}}$ 构造 $\boldsymbol{D},\boldsymbol{V}$
5	计算 $\boldsymbol{AVD}^{-1}$ 的 SVD，令 $\boldsymbol{w}$ 为右奇异向量
6	返回 $\boldsymbol{v} = \boldsymbol{VD}^{-1}\boldsymbol{w}$

该算法的实质是，在 $\boldsymbol{V}$ 的基中选择一个解。矩阵 $\boldsymbol{D}^{-1}$ 为 $\boldsymbol{V}$ 的列赋予权值，因此，接近于纯平解的向量获得较高的价值。与 $\boldsymbol{w}$ 关联的奇异值是式(5.30)的极小值。对于一个含较低噪声的系统，期望其值逼近于零，除非那个真实表面就是一个平凡解。

5.4.3 透视投影系统的求解

由前可知，对于正投影系统，有 $\boldsymbol{A}\boldsymbol{v}=0$；对于透视投影系统，有 $\boldsymbol{A}_f\boldsymbol{v}_f=0$。

定理 5.2(正投影和透视投影的等价性)

设 D 是一个已标签化的图像，p 为像平面前方的任一个 3D 点。那么，D 为一个多面体关于以 p 为中心的透视投影，当且仅当 D 为该多面体的正投影。

证明：

在不改变 D 可视性的情况下，通过变换 (x,y,z) 坐标系统，并调整缩放比例，使得 p 与原点一致，且使得像平面位于 $z=1$ 的表面上。由此可得到求证结论。

因此，场景的可视性既不依赖于观察者所站位置，也不依赖于是透视投影还是正投影。

类似的，正投影系统具有一个精确非平凡解，当且仅当具有同样交叉点的透视投影系统也有一个精确解。

□

这意味着，给定一个无噪系统，可以检验：曲线在一种非平凡构型(透视)中是否是真正平坦的，即存在精确解，而无须知道相机的焦距，根据正投影结论即可验证。因为如果是正投影，总能检查是否存在一个解。同时可将这个结论推广到含噪系统。

在透视情况下，用 f 去除矩阵 $\boldsymbol{A},\boldsymbol{Z},\boldsymbol{P}$ 中的元素 x_{ij},y_{ij},x_i,y_i，得到

$\boldsymbol{A}_f, \boldsymbol{Z}_f, \boldsymbol{P}_f$。再计算类似于式(5.23)中的$\boldsymbol{C}_f$。而在正投影情况下，矩阵$\boldsymbol{C}_f$执行一次平面拟合。

但在这里不再去为 3D 点集(X_i, Y_i, Z_i)拟合一个平面，在代数上更方便的方法是去定义$\boldsymbol{C}_f$，用于度量变换后的点集，即

$$(\frac{X_i}{Z_i}=\frac{x_i}{f}, \frac{Y_i}{Z_i}=\frac{y_i}{f}, \frac{1}{Z_i}) \tag{5.32}$$

的平面性偏差。这样，如果变换后的点在下述的一个平面上，则对应的 3D 点在平面$aX_i+bY_i+cZ_i=1$上。

$$a\frac{X_i}{Z_i}+b\frac{Y_i}{Z_i}+c=\frac{1}{Z_i}$$

根据这些定义，需要求解式(5.33)。

$$\arg\min_{\boldsymbol{v}_f}\left\|\boldsymbol{A}_f\boldsymbol{v}_f\right\| \quad \text{s.t.} \quad \left\|\boldsymbol{C}_f\boldsymbol{v}_f\right\|=1 \tag{5.33}$$

对正如正投影系统的任何解向量为$\boldsymbol{v}=(a_1,\cdots,a_N,b_1,\cdots,b_N,d_1,\cdots,d_N)^{\mathrm{T}}$，分配一个解$\boldsymbol{v}_f$给透视系统。

$$\boldsymbol{v}_f=(fa_1,\cdots,fa_N,fb_1,\cdots,fb_N,d_1,\cdots,d_N)^{\mathrm{T}} \tag{5.34}$$

通过直接计算，得到

$$\begin{cases}\boldsymbol{A}_f\boldsymbol{v}_f=\boldsymbol{A}\boldsymbol{v}\\ \boldsymbol{Z}_f\boldsymbol{v}_f=\boldsymbol{Z}\boldsymbol{v}\\ \boldsymbol{P}_f\boldsymbol{P}_f^{+}=\boldsymbol{P}\boldsymbol{P}^{+}\\ \boldsymbol{C}_f\boldsymbol{v}_f=\boldsymbol{C}\boldsymbol{v}\end{cases} \tag{5.35}$$

这样，根据式(5.30)和式(5.33)得到的极小奇异值是相同的。需

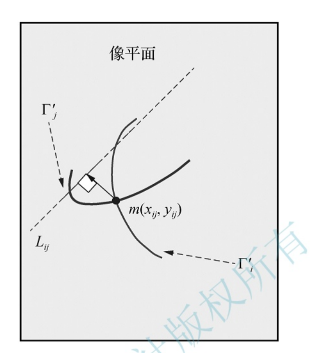
要说明的是，如果 $\boldsymbol{v}$ 垂直于基本平凡子空间，则结果就是 $\boldsymbol{v}_f$。不过，如果 $\boldsymbol{v}$ 垂直于 $\boldsymbol{C}$ 的零向量空间，则不能确保 $\boldsymbol{v}_f$ 垂直于 $\boldsymbol{C}_f$ 的零向量空间。

但是，选择一个垂直于平凡子空间的解的做法对于透视投影不太适用，因为透视投影可能放置一些可见点在相机的后面，这违背了 Cheirality 约束。所谓 Cheirality 约束，是指图像中的任何点必须位于生成该图像的相机的前面。根据这个思想，可以从类似仿射变换之类的两视图中确定场景结构。因此，在某些情况下，解必须含有一个平凡分量。

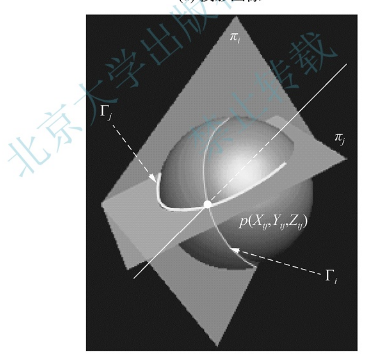
5.5 改善精确度

实验表明，上述方法比简单的线性方法更稳定、更可靠、适应性更强。但是，有时候这种方法也可能遭遇失败，主要原因是，虽然曲线交叉点的投影在图像中呈近似线性排列，而实际交叉点在 3D 中却不是线性的。

对于含噪系统，可以构造另一种不同的误差模型，即假设图像中被观察到的交叉点被独立同分布的各向同性高斯噪声所污染。通过极小化观察到的交叉点 (x_{ij}, y_{ij}) 与所估计的平面交叉线 L_{ij} 的投影线之间的距离的平方和，最终实现最大似然估计，如图 5.23 所示。



(a) 投影图像



(b) 3D空间构型

图 5.23　交叉点与交叉线的投影图像

对于正投影系统，最大似然估计为

$$\sum_{ij}\frac{[(a_i-a_j)x_{ij}+(b_i-b_j)y_{ij}+(d_i-d_j)]^2}{(a_i-a_j)^2+(b_i-b_j)^2} \quad \text{s.t.} \ \boldsymbol{v}\perp \text{Null}(\boldsymbol{C}) \tag{5.36}$$

对于透视投影系统，最大似然估计为

$$\sum_{ij}\frac{[(a_i-a_j)x_{ij}/f+(b_i-b_j)y_{ij}/f+(c_i-c_j)]^2}{(a_i-a_j)^2+(b_i-b_j)^2} \quad \text{s.t.} \ \boldsymbol{v}\perp \text{Null}(\boldsymbol{C}_f) \tag{5.37}$$

在式(5.36)中，分子与式(5.10)一样。对于不同的向量$\boldsymbol{v}$，分数呈现同构。当加入基本平凡分量向量$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$，式(5.36)将保持不变。Kawasaki 等人[178]沿着式(5.36)中的直线进行优化，而不需要$\boldsymbol{v}\perp \text{Null}(\boldsymbol{C})$约束条件。若没有这个约束条件，优化过程可以自由增加无边界的平凡分量，并收敛于$0/0$。但在数值上这会招致较大麻烦。有很多数值方法可用于式(5.36)，其中有些方法是专门设计用来极小化齐次有理函数的[179,180]。这些数值方法典型地收敛于一个局部极小值，但依赖于初始假定值。

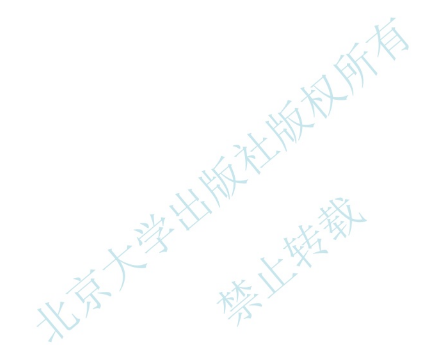
当曲线方程组是负定的，平面性约束可与附加的深度信息相结合。例如，Lipson 和 Shpitalni[166]最大化度量诸如拐角正交性、线段垂直性等。另一个例子是 SOR(Surfaces Of Revolution)[164]，这是含有两个附加约束的 SHGC。Shimshoni 和 Ponce[174]及 Shimodaira[119]使用明

暗约束信息。这些工作需要非线性优化，可能导致高阶运算。相比之下，Sugihara[156]提出要简化搜索思想，即首先解决线性约束以得到一个包含解的低维线性子空间，然后搜索缩放因子和平凡子空间系数，以优化附加的目标。这种方法的好处是，非线性搜索局限在非常低维的空间。如果忽略全局深度歧义性(即 $\boldsymbol{v}_3$)，GBR 歧义性的 3D 空间即使被稠密采样，也足以逼近非线性目标函数的解。这样有时候可帮助拓展搜索空间，使用附加的相当于较小奇异值的 $\boldsymbol{C}$ 的奇异向量。

5.6　本 章 小 结

共平面性是图像几何的一个重要特性，包含线面的平行性和点的共线性。人类视觉系统在感知立体场景时自觉利用了共平面性假设，一定程度上可弥补投影过程中丢失的景深信息。已有的共平面性约束方法存在两个极端，或者考察全部的自由变量，或者仅针对共面点深度的稀疏约束，都难以直接用于单图像的 3D 表面重构。基于传统方法的这种局限性，本章提出一种基于交叉平坦曲线的共平面约束的表面重构方法，即对于不平行于投影方向的任一个平面，根据其所含曲线与另一个平面中某曲线的交叉构型构造一个线性系统，当这些交叉曲线位于拟求解表面时，可获得该系统的精确解空间。对于含噪系统，增加 $\|\boldsymbol{C}\boldsymbol{v}\|=1$ 约束条件，使用 SVD 法得到逼近解。利用正投影和透视

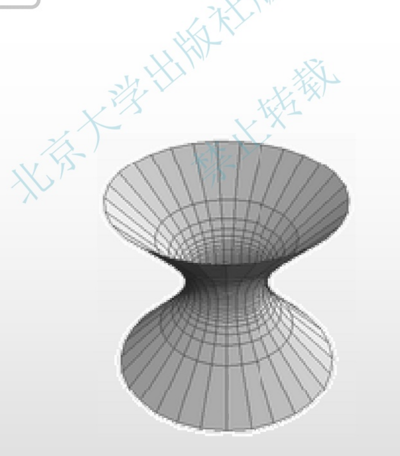
投影的等价性，可将透视投影转化为正投影，从而将基于平坦交叉曲线的 3D 表面重构问题统一化，并分别提出了改善解精确度的最大似然估计函数。实验证明，这种方法大大改良了单图像 3D 表面重构的健壮性，可用于单视图模型化、完全未标定结构光等情况。





第 6 章

总结与展望



计算机视觉的终极目标是从所“看”到的场景中识别出感兴趣的对象，即基于 2D 图像重构 3D 模型。但是，由于成像过程丢失了景深信息或因遮挡而遗漏了部分有用信息，导致不同形状的 3D 物体投影在图像平面上产生了相同的图像，这就是投影歧义性。基于单幅图像的 3D 重构是计算机视觉领域极具挑战性的研究热点，由于其所能提供的信息非常有限，因此是一个经典的非适定难题，需要借助一些关于所研究对象的先验知识或经验常识，如几何形状的平坦性、平行性及垂直性，表面纹理的畸变性，局部表面的平滑性、明暗渐变性、散焦或聚焦性，增加人机交互等，以将其驯服为适定问题。

6.1 本书工作总结

目前常用的 3D 重构方法主要有 SFS、SFT 等，其对先验知识的假定大多是试探性的，仅适合于特定领域、特殊场合或特定问题，如对称物体(人脸、建筑物)等。本书针对单幅图像 3D 表面重构中广泛存在的歧义性问题，将约束条件归纳为线性、离散和二次约束三种，并在以下三个方面进行了探索。

(1) 根据表面可积性约束歧义性，将 SFS 规划为含二次约束的多项式系统，利用同伦分析法、优化线搜索法和半定规划法求解该多项式，并设计一种新的平滑项，使 SFS 摆脱了对初始解的依赖性，优化了迭代过程，对非凸项实现了统一的 SDP 凸松弛处理。这种将单图像的 3D 表面重构问题归入到一个成熟的理论框架的一种尝试，对 SFT、

SFM 等重构方法的进一步研究都具有重要的参考价值。实验结果表明，这种处理方案大大改善了 SFS 方法的高效性和普适性。

(2) 将纹理图案中广泛存在的二进制歧义性问题归结为一个离散约束问题，将表面表示为一个张量积基函数的线性组合，进而定义一个二次目标函数来度量表面平滑度和约束满足度，通过消除连续变量产生一个仅含离散变量的二次目标函数，再使用 SDP 凸松弛法，将离散变量嵌入一个连续高维空间，通过基于分裂惯性和角扫描的圆整过程将离散变量归约到一个低维空间，最后通过最大割算法获得较理想的离散判定结果。上述方案比传统的椭球法高效且成本低($O(n^2)$)，还可进一步使用 K-L 算法改良搜索过程，消除错误割点，大大提高解的精确度。

(3) 将图像中诸如 in/out 倒置的几何歧义性规划为共平面线性约束条件，提出一种基于交叉平坦曲线的 3D 表面重构方法，即对于不平行于投影方向的任一个平面，根据其所含曲线与另一个平面中某曲线的交叉构型构造一个线性系统，当这些交叉曲线位于拟求表面时，可获得该系统的精确解空间。对于含噪系统，增加$\|\boldsymbol{Cv}\|=1$约束条件，使用 SVD 法得到逼近解。利用正投影和透视投影的等价性，可将透视投影转化为正投影，从而将基于平坦交叉曲线的 3D 表面重构模式统一化，并分别设计了改善解精确度的最大似然估计函数。实验证明，这种方法大大提高了单图像 3D 表面重构的健壮性和高效性，可用于完全未标定结构光等真实场景。

6.2　进一步研究展望

基于单幅图像的 3D 表面重构是一个经典的约束不足难题，虽然经过大量研究人员的多年不懈努力取得了可喜的成就，但在基础理论方面仍然存在很大的探索空间，这些年涌现出的各种解决方案或应用技术也大多侧重于解决特定领域的特定问题。本书在此方面做了大量的研究，取得的研究成果从一个侧面推动了该领域的进展，同时发现在实时性重构等方面还存在一些问题需要进行更细致深入的探索。

(1) 现有的单图像重构方案未能实现对实时噪声的有效抑制，尤其是因摄像设备因素、恶劣环境等产生的随机噪声没有统一、有效的数学模型，针对特定场合设计的噪声模型的实际效果也非常有限，微小的噪声干扰也可能导致多重歧义解。这种情况主要是因为缺乏有效的数学模型，因而极大地限制了此类技术在对实时性要求较高的场合中的推广应用，诸如智能交通中的压黄线违章检测、远程视频会诊、网络游戏等。今后将继续关注噪声表示方法的理论探索。

(2) 目前的 SFT 方法大多针对有规则的确定性纹理，如树叶、建筑墙面等。但这种纹理毕竟很特殊，对于室外天然草坪、山坡、海平面等统计性纹理，常规 SFT 方法根本不适用，主要因为对室外自然场景中的无规则纹理特征没有统一、有效的处理方式，需要寻求更合理的数学模型。本书与其他方法一样，对于纹理歧义性的多样性、可叠加性等考虑不足，今后将进一步寻求理论上的突破。

(3) SFS 方法一般假设为单光源，单光源模型比较简单，而实际情况可能是多种类、多个光源的交互作用，如夜间的街景，存在多个路灯、车灯，以及月光等，这些光源的随机组合形成了实际的光照效果。这些偶发光源将切实影响单光源的假定前提，本书与其他方法一样没有对此深入研究，今后将致力于光源随机组合方面的研究，希望能找出一种有效的光源组合策略和解转换方法，能将单光源下的研究成果无缝推广到多种类、多地缘的复杂源环境下。

可以预见的是，如果能从理论上解决上述问题，或在技术或具体的方法上得以长足改进，必将极大地推动基于图像的 3D 重构技术在智能交通、远程诊疗、航空航天、真人游戏等方面的广泛应用，进而带动相关学科的协调发展。

参 考 文 献

[1] MARR D, POGGIO T, 1979. A Computational Theory of Human Stereo Vision[J]. Royal Society of London, 204(B): 301-328.

[2] MARR D, 1982. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information[M]. New York: W.H. Freeman & Company:501-524.

[3] DYER C R, 2001. Volumetric Scene Reconstruction from Multiple Views[M]//. Foundations of Image Understanding. Boston: Kluwer: 469-489.

[4] SLABAUGH G, CULBERTSON B, MALZBENDER T, et al, 2001. Survey of Methods for Volumetric Scene Reconstruction from Photographs[C]. Joint IEEE TCVG and Eurographics Workshop in Volume Graphics.

[5] LAZEBNIK S, FURUKAWA Y, PONCE J, 2007. Projective Visual Hulls[J]. International Journal of Computer Vision.

[6] AGANJ E, PONS J P, et al, 2007. Spatio-Temporal Shape from Silhouette Using Four Dimensional Delaunay Meshing[C]. IEEE International Conference on Computer Vision.

[7] HAN F, ZHU S C, 2003. Bayesian Reconstruction of 3D Shapes and Scenes from a Single Image[C]. The First IEEE International Workshop on Higher-Level Knowledge in 3D Modeling and Motion Analysis.

[8] RAMANAN D, 2006. Learning to Parse Images of Articulated Objects[C]. Advances in Neural Information Processing Systems.

[9] BOWDEN R, MITCHELL T A, SARHADI M, 1998. Reconstructing 3D Pose and

Motion from a Single Camera View[C]. British Machine Vision Conference.

[10] HOWE N R, LEVENTON M E, Freeman W T, 2000. Bayesian Reconstruction of 3D Human Motion from Single-Camera Video[C]. Advances in Neural Information Processing Systems.

[11] SIGAL L, BLACK M J, 2006. Predicting 3D People from 2D Pictures[C]. IV Conference on Articulated Motion and Deformable Objects.

[12] Longuet-Higgins H, 1981. A Computer algorithm for Reconstructing a Scene From Two Projections[J]. Nature, 293:133-135.

[13] HORRY H, ANJYO K Arai K, 1997. Tour into the Picture: Using a Spidery Mesh Interface to Make Animation from a Single Image**[C].** ACM SIGGRAPH: 225-232.

[14] CRIMINISI A, REID I, ZISSERMAN A, 1999. Single View Metrology[C]. International Conference on Computer Vision:434-442.

[15] SHUM H Y, HAN M, SZELISKI R, 1998. Interactive Construction of 3D Models from Panoramic Mosaics[C]. CVPR: 427-433.

[16] DEBEVEC P, TAYLOR C, Malik J, 1996. Facade: Modeling and Rendering Architecture from Photographs[C]. ACM SIGGRAPH : 11-20.

[17] COORG S, TELLER S, 1999. Extracting Textured Vertical Facades from Controlled Close-range Imagery[C]. CVPR.

[18] COLLINS R T, 1996. A Space-sweep Approach to True Multi-image Matching[C]. CVPR.

[19] Gallup D, Frahm J M, Mordohai P, et al, 2007. Real-Time Plane-Sweeping Stereo with Multiple Sweeping Directions[C]. CVPR: 1-8.

[20] YANG A Y, HUANG K, RAO S, et al, 2005. Symmetry-Based 3-D

Reconstruction from Perspective Images[J]. Computer Vision and Image Understanding, 99(2): 210-240.

[21] DEBEVEC P E, TAYLOR C J, MALIK J, 1996. Modeling and Rendering Architecture from Photographs: A Hybrid Geometry and Image Based Approach[C]. SIGGRAPH.

[22] GROSSMANN E, ORTIN D, SANTOS J, 2002. Victor Single and Multi-view Reconstruction of Structured Scenes[C]. ACCV.

[23] LEIBOWITZ D, CRIMINISI A, ZISSERMAN A, 1999. Creating Architectural Models from Images[C]. Eurographics: 39-50.

[24] CIPOLLA R, ROBERTSON D, 1999. 3D Models of Architectural Scenes from Uncalibrated Images and Vanishing Points[C]. 10th IAPR: 824-829.

[25] Strum P F, MAYBANK S J, 1999. A Method for Ineractive 3D Reconstruction of Piecewise Planar Objects from Single Views[C]. BMVC: 265-274.

[26] WILCZKOWIAK M, BOYER E, STRUM P, 2001. Camera Calibration and 3D Reconstruction from Single Images Using Parallelepipeds[C]. ICCV: 142-148.

[27] SCHARSTEIN D, SZELISKI R, 2002. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms[J]. International Journal of Computer Vision.

[28] ANGUELOV D, SRINIVASAN P, KOLLER D, et al, 2005. SCAPE: Shape Completion and Animation of People[C]. SIGGRAPH: 408-416.

[29] AGARWAL A, TRIGGS B, 2004. Learning to Track 3D Human Motion from Silhouettes[C]. 21st ICML.

[30] SNOW D, VIOLA P, ZABIH R,2000. Exact Voxel Occupancy with Graph Cuts[C]. CVPR.

[31] BALAN A O, SIGAL L, BLACK M J, 2007. Detailed Human Shape and Pose from Images[C], CVPR.

[32] SEEMANN E, LEIBE B, SCHIELE B, 2006. Multi-aspect Detection of Articulated Objects[C]. CVPR.

[33] HOIEM D, EFROS A, HEBERT M, 2005. Automatic Photo Pop-up[C]. SIGGRAPH: 577-584.

[34] PRASAD M, ZISSERMAN A, FITZGIBBON A, 2006. Single View Reconstruction of Curved Surfaces[C]. CVPR:1345-1354.

[35] TERZOPOULOS D, WITKIN A, KASS M, 1987. Symmetry-Seeking Models and 3D Object Reconstruction[J] IJCV, 1: 211-221.

[36] BLANZ V, VETTER T, 1999. A, Morphable Model for the Synthesis of 3D Faces[C]. SIGGRAPH: 187-194.

[37] SIGAL P, BĂLAN A, BLACK M, 2007. Combined Discriminative and Generative Articulated Pose and Non-rigid Shape Estimation[C]. Advances in Neural Information Processing System.

[38] Horn B P K, 1970. Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View[D]. Boston: MIT.

[39] SHIMSHONI I, MOSES Y, LINDENBAUM M, 2000. Shape Reconstruction of 3D Bilaterally Symmetric Surfaces[J]. IJCV 39(2): 97-110.

[40] ZHAO W, CHELLAPPA R, 1999. Robust Face Recognition Using Symmetric Shape-from-Shading[R]. Technical Report CARTR-919, Center for Automation Research, University of Maryland.

[41] ZHAO W, CHELLAPPA R, 2000. Illumination-Insensitive Face Recognition Using Symmetric Shape-from-Shading[C]. CVPR: 286-293.

[42] RAMACHANDRAN V S, 1988. Perceiving Shape from Shading[J]. Scientific American, 159: 76-83.

[43] BICHSEL M, PENTLAND A P, 1992. A Simple Algorithm for Shape from Shading[C]. CVPR: 459-465.

[44] HATZITHEODOROU M, 1989. The Derivation of 3-D Surface Shape from Shadows[C]. Image Understanding Workshop: 1012-1020.

[45] GIBSON J J, 1950. The Perception of the Visual World[M]. Boston: Houghton Mifflin.

[46] GÅRDING J, 1992. Shape from Texture for Smooth Curved Surfaces in Perspective Projection[J]. JMIV, 2: 327-350.

[47] HWANG W L, LU C S, CHUNG P C, 1998. Shape from Texture: Estimation of Planar Surface Orientation through the Ridge Surfaces of Continuous Wavelet Transform[J]. IEEE Transactions on Image Processing, 7: 773-780.

[48] KANATANI K, CHOU T C, 1989. Shape from Texture: General Principle[J]. Artificial Intelligence, 38: 1-48.

[49] FORSYTH D A, 2001. Shape from Texture and Integrability[C]. 8th ICCV, 2:447 - 452.

[50] KOENDERINK J J, DOORN A J, Kappers A M L, et al, 2001. Ambiguity and the "Mental Eye" in Pictorial Relief[J]. Perception, 30(4):431-448.

[51] LANGER M S, BÜLTHOFF H H, 1999. Perception of Shape from Shading on a Cloudy Day[R]. Technical Report 73, Max Planck Institute for Biological Cybernetics.

[52] NAKAYAMA K, HE Z J, SHIMOJO S, 1995. Visual Surface Representation: A Critical Link Between Lower-Level and Higher-Level Vision[M]//An Invitation

to Cognitive Science. 2nd ed. Cambridge: MIT Press:1-70.

[53] PALMER S E, 1999. Vision Science: Photons to Phenomenology[M]. Cambridge: MIT Press.

[54] SHEPARD R N, METZLER J, 1971. Mental Rotation of Three-Dimensional Objects[J]. Science,171(3972):701-703.

[55] TYLER C W, 2006. Spatial Form as Inherently Three Dimensional[M]//Seeing Spatial Form. Oxford: Oxford University Press: 67-88.

[56] ROBERTS L G, 1963. Machine Perception of Three-Dimensional Solids[D]. Boston: MIT.

[57] CHEN H F, BELHUMEUR P N, JACOBS D W, 2000. In Search of Illumination Invariants[C]. CVPR: 254-261.

[58] NECKER L A, 1832. Observations on Some Remarkable Optical Phaenomena Seen in Switzerland; and on an Optical Phaenomenon Which Occurs on Viewing a Figure of a Crystal or Geometrical Solid[J]. London and Edinburgh Philosophical Magazine and Journal of Science, 1 (5): 329-337.

[59] KANATANI K, 1988. Constraints on Length and Angle[J]. CVGIP, 41(1):28-42.

[60] VARLEY P A C, MARTIN R R, 2002. Estimating Depth from Line Drawings[C]. 7th ACM Symposium on Solid Modeling and Applications: 180-191.

[61] KANATANI K, 2008. Latest Progress of 3-D Reconstruction from Multiple Camera Images[J]. Robotics Research Trends: 33-75.

[62] COLE J D, 1992. Perturbation Methods in Applied Mathematics[M]. Waltham: Blaisdell Publishing Company.

[63] KEVORKIAN J, COLE J D, 1995. Multiple Scales and Singular Perturbation

Methods[M]. Berlin: Springer-Verlag.

[64] LINDSTEDT A, 1882. Under Die Integration Einer Fur Die Storungstheorie wichtigen differentialgleichungAstron Nach 103:211-222.

[65] MURDOCK J A, 1991. Perturbations-Theory and Methods[M]. New York: John Wiley & Sons.

[66] NAYFEH A H, 2000. Perturbation Methods[M]. New York: John Wiley & Sons.

[67] SEN S, 1983. Topology and Geometry for Physicists[M]. Florida: Academic Press.

[68] LYAPUNOV A M, 1992. GENERAL Problem on Stability of Motion translation[M]. London: Taylor & Francis.

[69] AWREJCEWICZ J, ANDRIANOV I V, MANEVITCH L I, 1998. Asymptotic Approaches in Nonlinear Dynamics[M]. Berlin: Springer-Verlag.

[70] KARMISHIN A V, ZHUKOV A T, KOLOSOV V G, 1990. Methods of Dynamics Calculation and Testing for Thin-walled Structures[M]. Moscow: Mashinostroyenie.

[71] ADOMIAN G, 1976. Nonlinear Stochastic Differential Equations[J]. J MATH ANAL APPL, 55: 441-452.

[72] ADOMIAN G, 1991. A Review of the Decomposition Method and Some Recent Results for Nonlinear Equations[J]. COMPUT MATH APPL, 21: 101-127.

[73] ADOMIAN G, 1994. Solving Frontier Problems of Physics: The Decomposition Method[M]. Boston: Kluwer Academic Publisher.

[74] ADOMIAN G, ADOMIAN G E, 1984. A Global Method for Solution of Complex Systems[J]. Math Model: 521-568.

[75] RACH R, 2008. A New Definition of Adomian Polymonial[J]. Kybernetes: 910-955.

[76] LIAO S J, 1992. The Proposed Homotopy Analysis Technique for the Solution of Nonlinear Problems[D]. Shanghai: Shanghai Jiao Tong University.

[77] HILTON P J, 1953. An Introduction to Homotopy Theory[M]. Cambridge: Cambridge University Press.

[78] LIAO S J, 1997. A Kind of Approximate Solution Technique Which Does Not Depend Upon Small Parameters (II) - An Application in Fluid Mechanics[J]. INT J NONLINMECH, 32(5): 815-822.

[79] LIAO S J, 2003. Beyond Perturbation: Introduction to the Homotopy Analysis Method[M]. Roca Raton: Chapman & Hall/ CRC Press.

[80] LIAO S J, 2012. The Homotopy Analysis Method in Nonlinear Differential Equations[M]. BeiJing: Higher Education Press.

[81] LIAO S J, 1999. An Explicit, Totally Analytic Approximation of Blasius Viscous Flow Problems[J]. INT J NONLIN MECH,34(4): 759-778.

[82] YABUSHITA K, YAMASHITA M, TSUBOI K, 2007. An Analytic Solution of Projectile Motion with the Quadratic Resistance Law Using the Homotopy Analysis Method[J]. J PHYS A - MATH THEORY, 40: 8403-8416.

[83] AKYILDIZ F T, VAJRAVELU K, 2008. Magnetohydrodynamic Flow of a Viscoelastic Fluid[J]. Physics Letter A, 372: 3380-3384.

[84] MARINCA V, HERISANU N, 2008. Application of Optimal Homotopy Asymptotic Method for Solving Nonlinear Equations Arising in Heat Transfer[J]. INT COMMUN HEAT MASS, 35:710-715.

[85] BHAT K, TWIGG C, HODGINS J, et al, 2003. Estimating Cloth Simulation

Parameters from Video[C]. SCA: 3751.

[86] COHEN L, COHEN I, 1993. Finite-element Methods for Active Contour Models and Balloons for 2-D and 3-D Images[J]. Trans PAMI, 15(11):1131-1147.

[87] MCINERNEY T, TERZOPOULOS D, 1993. A finite Element Model for 3D Shape Reconstruction and Nonrigid Motion Tracking[C]. ICCV: 518-523.

[88] METAXAS D, TERZOPOULOS D, 1993. Constrained Deformable Superquadrics and Nonrigid Motion Tracking[J]. Trans PAMI, 15(6):580-591.

[89] TSAP V, GOLDGOF D B, SARKAR S, 2000. Nonrigid Motion Analysis Based on Dynamic refinement of Finite Element Models[J]. Trans PAMI, 22(5):526-543.

[90] COOTES T, EDWARDS G, TAYLOR C, 1998. Active Appearance Models[C]. ECCV: 484-498.

[91] FUA P, LECLERC Y G, 1995. Object-centered Surface Reconstruction: Combining Multi-image Stereo and Shading[J]. Int Journal Comp Vision, 16:35-56.

[92] MATTHEWS I, BAKER S, 2004. Active Appearance Models Revisited[J]. Int Journal Comp Vision, 60:135-164.

[93] PERRIOLLAT M, HARTLEY R, BARTOLI A, 2008. Monocular Template-Based Reconstruction of Inextensible Surfaces[C]. BMVC.

[94] HORN B P K, SZELISKI R S, YUILLE A L, 1993. Impossible Shaded Images[J]. PAMI,15(2):166-170.

[95] OLIENSIS J, 1991. Uniqueness in Shape from Shading[J]. IJCV, 6(2):75-104.

[96] SAXBERG B V H, 1989. A Modern Differential Geometric Approach to Shape from Shading[D]. Boston: MIT.

[97] OKATANI T, DEGUCHI K, 2000. Closed form Solution of Local Shape from Shading at Critical Points[J]. IJCV, 40(2):169-178.

[98] DEIFT P, SYLVESTER J, 1981. Some Remarks on the Shape-from-Shading Problem in Computer Vision[J]. J.Math Anal Appl, 84(1):235-248.

[99] ZHANG R, TSAI P S, Cryer J E, et al, 1999. Shape from Shading: A Survey[J]. PAMI, 21(8):690-706.

[100] LEE K M, KUO C C J, 1994. Shape from Shading with Perspective Projection[J]. CVGIP:Image Understanding, 59(2):202-212.

[101] AHMED A H, FARAG A A, 2006. A New Formulation for Shape from Shading for Nonlambertian Surfaces[C]. CVPR: 1817-1824.

[102] KAO C Y, OSHER S, QIAN J, 2004. Lax-Friedrichs Sweeping Scheme for Static Hamiltonjacobi Equations[J]. Journal of Computational Physics, 196(1):367-391.

[103] ROUY E, TOURIN A, 1992. A Viscosity Solutions Approach to Shape-from-Shading SIAM[J]. Journal on Numerical Analysis, 29(3):867-884.

[104] KIMMEL K, SETHIAN J A, 2001. Optimal Algorithm for Shape from Shading and Path Planning[J]. Journal of Mathematical Imaging and Vision, 14(3):237-244.

[105] OLIENSIS J, DUPUIS P, 1991. Direct Method for Reconstructing Shape from Shading[J]. SPIE, 1570:116-128.

[106] PRADOS E, 2004. Application of the Theory of the Viscosity Solutions to the Shape from Shading Problem[D]. Sophia Antipolis: INRIA.

[107] PENNA M A, 1989. A shape from Shading Analysis for a Single Perspective Image of a Polyhedron[J]. PAMI, 11(6):545-554.

[108] PRADOS E, FAUGERAS O, 2005. A Generic and Provably Convergent Shape-from-Shading Method for Orthographic and Pinhole Cameras[J]. IJCV, 65(1-2):97-125.

[109] TANKUS A, SOCHEN N, YESHURUN Y, 2005. Shape-from-Shading under Perspective Projection[J]. IJCV, 63(1):21-43.

[110] YUEN S Y, TSUI Y Y, CHOW C K, 2007. A Fast Marching Formulation of Perspective Shape from Shading under Frontal Illumination[J]. Pattern Recognition Letters, 28(7):806-824.

[111] ECKER A, JEPSON A D, KUTULAKOS K N, 2008. Semidefinite Programming Heuristics for Surface Reconstruction Ambiguities[C]. ECCV: 12-14.

[112] SALZMANN M, MORENO-NOGUER N, LEPETIT V, et al, 2008. Closed-form Solution to Non-rigid 3D Surface Registration[C]. ECCV: 581-594.

[113] WHITE R, FORSYTH D, 2006. Combining Cues: Shape from Shading and Texture[C]. CVPR: 1809-1816.

[114] FAUGERAS O D, LUONG Q T, MAYBANK S J, 1992. Camera Self-calibration: Theory and Experiments[C]. CVPR: 321-334.

[115] BENNETT B M, HOFFMAN D D, NICOLA J E, et al, 1989. Inferring Structure from Motion: A Homotopy Algorithm[C]. Workshop on Visual Motion: 238-245.

[116] WEINBERGER K Q, SAUL L K, 2006. Unsupervised Learning of Image Manifolds by Semidefinite Programming[J]. IJCV, 70(1):77-90.

[117] LEE K M, KUO C C J, 1993. Shape from Shading with a Linear Triangular Element Surface Model[J]. PAMI, 15(8):815-822.

[118] YANG J, OHNISHI N, ZHANG D, 1997. Determining a Polyhedral Shape Using Interreflections[C]. CVPR: 110-115.

[119] SHIMODAIRA H, 2006. A Shape-from-Shading Method of Polyhedral Objects Using Prior Information[J]. PAMI, 28(4):612-624.

[120] SOMMESE A J, WAMPLER C W, 2005. The Numerical Solution of Systems of Polynomials, Arising in Engineering and Science[M]. Singapore: World Scientific Publishing Company.

[121] LUO Z Q, ZHANG S, 2010. A Semidefinite Relaxation Scheme for Multivariate Quartic Polynomial Optimization with Quadratic Constraints[J]. SIOPT, 20(4):1716-1736.

[122] QI L, WAN Z, YANG Y F, 2005. Global Minimization of Normal Quartic Polynomials Based on Global Descent Directions[J]. SIOPT, 15(1):275-302.

[123] SZELISKI R, 1990. Fast Shape from Shading[C]. ECCV: 359-368.

[124] GÜltekin A, GÖKMEN M, 1998. Adaptive Shape from Shading[J]. Turkish Journal of Electrical Engineering and Computer Sciences,6(2):61-73.

[125] HUANG R, SMITH W A P, 2009. A Shape-from-Shading Framework for Satisfying Datacloseness and Structure-Preserving Smoothness Constraints[C]. BMVC.

[126] BISWAS P, AGHAJAN H, YE Y, 2005. Semidefinite Programming Algorithms for Sensor Network Localization Using Angle Information[C]. Asilomar Conference on Signals, Systems and Computers: 220-224.

[127] BISWAS P, LIAN T C, WANG T C, 2006. Semidefinite Programming Based Algorithms for Sensor Network Localization[J]. ACM Transactions on Sensor Networks (TOSN),2(2):188-220.

[128] FAZEL M, HINDI H, BOYD S P, 2003. Log-det Heuristic for Matrix Rank Minimization with Applications to Hankel and Euclidean Distance Matrices[C]. American Control Conference, 3: 2156-2162.

[129] WAKI H, KIM S, KOJIMA M, et al, 2006. Sums of Squares and Semidefinite Program Relaxations for Polynomial Optimization Problems with Structured Sparsity[J]. SIAM Journal on Optimization, 17(1):218-242.

[130] NIE J, MEVISSEN M, KOJIMA M, et al, 2008. Solving Partial Differential Equations via Sparse SDP Relaxations[J]. Pacific Journal of Optimization, 4(2):213-241.

[131] HENRION D, LASSERRE J B, 2005. Detecting Global Optimality and Extracting Solutions in GloptiPoly, In Positive Polynomials in Control[M]. Berlin: Springer-Verlag: 293-310.

[132] FORSYTH D A, 2002. Shape from Texture without Boundaries[C]. ECCV: 225-239.

[133] LOBAY A, FORSYTH D A, 2004. Recovering Shape and Irradiance Maps from Rich Dense Texton Fields[C]. CVPR.

[134] LOH A, HARTLEY R, 2005. Shape from Non-homogeneous, Nonstationary, Anisotropic, Perspective Texture[C]. BMVC.

[135] CHANG J Y, LEE K M, LEE S U, 2008. Shape from Shading Using Graph Cuts[J]. Pattern Recognition, 41(12):3749-3757.

[136] NAITO S, ROSENFELD A, 1988. Shape from Random Planar Features[J]. CVGIP, 42(3):345-370.

[137] KOENDERINK J J, DOORN A J. Shape from Chebyshev Nets[C]. ECCV: 215-225.

[138] ARORA S, BERGER E, HAZAN E, et al, 2005. On Non-approximability for Quadratic Programs[C]. FOCS: 206-215.

[139] BERGER M J, BOKHARI S H A, 1987. Partitioning Strategy for Non-uniform Problems Across Multiprocessors[J]. IEEE Transactions on Computers, C-36:570-580.

[140] LELAND R, HENDRICKSON B, 1994. An Empirical Study of Static Load Balancing Algorithms[C]. Scalable High Performance Computing Conference: 682-685.

[141] CHAN T F, GILBERT J R, TENG S H, 1995. Geometric Spectral Partitioning[R]. Technical Report CSL-94-15, Xerox PARC.

[142] FJÄLLSTRÖM P O, 1998. Algorithms for Graph Partitioning: A Survey Linkoping Electronic[J]. Computer and Information Science, 3(10).

[143] KERNIGHAN B W, LIN S, 1970. An Efficient Heuristic Procedure for Partitioning Graphs[J]. The Bell System Technical Journal, 49(1):291-307.

[144] KARYPIS G, KUMAR V, 1998. A Fast and High Quality Multilevel Scheme for Partitioning Irregular Graphs[J]. SIAM Journal on Scientific Computing, 20(1):359-392.

[145] STEVENS K A, 1981. The Visual Interpretation of Surface Contours[J]. Artificial Intelligence, 17(1-3):47-73.

[146] TODD J T, REICHEL F D, 1990. Visual Perception of Smoothly Curved Surfaces from Doubleprojected Contour Patterns[J]. Journal of Experimental Psychology: Human Perception and Performance, 16(3):665-674.

[147] PIZLO Z, LI Y, FRANCIS G, 2005. A New Look at Binocular Stereopsis[J]. Vision Research, 45(17):2244-2255.

[148] GUZMAN G, 1968. Arenas Computer Recognition of Three-Dimensional Objects in a Visual Scene[D]. Boston: MIT.

[149] HUFFMAN D A, 1971. Impossible Objects as Non-sense Sentences[J]. Machine Intelligence, 6:295-323.

[150] CLOWES M B, 1971. On Seeing Things[J]. Artificial Intelligence, 2(1):79-116.

[151] KIROUSIS L, PAPADIMTRIOU C H, 1988. The Complexity of Recognizing Polyhedral Scenes[J].Computer and System Sciences, 37(1):14-38.

[152] WALTZ D L, 1975. Understanding Line Drawings of Scenes with Shadows[C]// The Psychology of Computer Vision. New York: McGraw-Hill: 19-91.

[153] MALIK D, 1987. Interpreting Line Drawings of Curved Objects[J]. IJCV, 1(1):73-103.

[154] VARLEY P A C, 2003. Automatic Creation of Boundary-Representation Models from Single Line Drawings[D]. Cardiff: Cardiff University.

[155] VARLEY P A C, MARTIN R R, Suzuki H, 2005. Frontal Geometry from Sketches of Engineering Objects: Is Line Labelling Necessary?[J]. Computer-Aided Design, 37(12):1285-1307.

[156] SUGIHARA K, 1986. Machine Interpretation of Line Drawings[M]. Cambridge: MIT Press.

[157] FALK G, 1972. Interpretation of Imperfect Line Data as a Three-Dimensional Scene[J]. Artificial Intelligence, 3:101-144.

[158] ROTHWELL C, STERN J, 1995. Understanding the Shape Properties of Trihedral Polyhedra[R]. Technical Report, INRIA.

[159] ROTHWELL C A, 1995. Object Recognition through Invariant Indexing[M]. Oxford: Oxford University Press.

[160] ROTHWELL C A, FORSYTH D A, ZISSERMAN A, et al, 1993. Extracting Projective Structure from Single Perspective Views of 3D Point Sets[C]. ICCV: 573-582.

[161] LAMDAN Y, WOLFSON H J, 1988. Geometric hashing: A General and Efficient Model-Based Recognition Scheme[C]. ICCV: 238-249.

[162] BURNS J B, WEISS R S, RISEMAN E M, 1992. The Non-existence of General-Case View-Invariants[C]//Geometric Invariance in Computer Vision. Cambridge: MIT Press: 120-131.

[163] REISS T H, 1993. Recognizing Planar Objects Using Invariant Image Features[M]. Berlin: Springer-Verlag.

[164] COLOMBO C, BIMBO A, PERNICI F, 2005. Metric 3D Reconstruction and Texture Acquisition of Surfaces of Revolution from a Single Uncalibrated View[J]. PAMI, 27(1):99-114.

[165] ULUPINAR F, NEVATIA R, 1995. Shape from Contour: Straight Homogeneous Generalized Cylinders and Constant Cross Section Generalized Cylinders[J]. PAMI, 17(2):120-135.

[166] LIPSON H, SHPITALNI M, 1996. Optimization-based Reconstruction of a 3D Object from a Single Freehand Line Drawing[J]. Computer-Aided Design,

28(8):651-663.

[167] SPARR G, 1992. Depth Computations from Polyhedral Images[J]. Image and Vision Computing, 10(10):683-688.

[168] HEYDEN A, 1996. On the Consistency of Line-Drawings, Obtained by Projections of Piecewise Planar Objects[J]. Journal of Mathematical Imaging and Vision, 6(4):393-412.

[169] WEYRICH T, DENG J, BARNES C, et al, 2007. Digital Bas-Relief from 3D Scenes[J]. ACM Transactions on Graphics (Proc. SIGGRAPH), 26(3).

[170] BELHUMEUR P N, KRIEGMAN D J, YUILLE A L, 1999. The Bas-Relief Ambiguity[J]. IJCV, 35(1):33-44.

[171] KOENDERINK J J, DOORN A J, 1991. Affine Structure from Motion[J]. JOSA A, 8(2):377-385.

[172] TSE P U, 2002. A Contour Propagation Account of Surface Filling-in and Volume Formation[J]. Psychological Review, 109(1):91-115.

[173] ROS L, THOMAS F, 2002. Overcoming Superstrictness in Line Drawing Interpretation[J]. PAMI, 24(4):456-466.

[174] SHIMSHONI I, PONCE J, 1997. Recovering the Shape of Polyhedra Using Line-Drawing Analysis and Complex Reflectance Models[J]. CVIU, 65(2): 296-310.

[175] GRIMSTEAD I J, MARTIN R R, 1995. Creating Solid Models from Single 2D Sketches[C]. 3rd ACM Symposium on Solid Modeling and Applications: 323-337.

[176] BOUGUET J Y, WEBER M, PERONA P, 1999. What Do Planar Shadows Tell Us About Scene Geometry?[C]. CVPR: 514-520.

[177] ULUPINAR F, NEVATIA R, 1993. Perception of 3-D Surfaces from 2-D Contours[J]. PAMI, 15(1):3-18.

[178] KAWASAKI H, FURUKAWA R, 2009. Shape Reconstruction and Camera Self-calibration Using Cast Shadows and Scene Geometries[J]. IJCV, 83(2):135-148.

[179] CHOJNACKI W, BROOKS M J, HENGEL A, et al, 2004. From FNS to HEIV:A link Between Two Vision Parameter Estimation Methods[J]. PAMI, 26(2):264-268.

[180] KANATANI K, 2008. Statistical Optimization for Geometric Fitting: Theoretical Accuracy Bound and High Order Error Analysis[J]. IJCV, 80(2):167-188.